A Quantitative Health Risk Assessment for the Kalamazoo River PCB site

Prepared for:

The Kalamazoo River Study Group

by:

Edmund Crouch, Ph.D., Michael Ames, Sc.D., and Laura Green, Ph.D., D.A.B.T.

June 1, 2001

Cambridge Environmental Inc

58 Charles Street Cambridge, Massachusetts 02141
617-225-0810 FAX: 617-225-0813 www.CambridgeEnvironmental.com

Contents

List of Tables

Summary

This report assesses risks to health for people exposed to polychlorinated biphenyl (PCB) contamination in and around the Kalamazoo River.[1] The people most highly exposed are anglers (and their families) who eat the fish they catch from the portions of the Kalamazoo River considered here.[2] Using U.S. Environmental Protection Agency criteria for acceptable risk to health, our risk assessment finds that further action to clean PCBs from the River may not be needed. That is, estimates of current and future risks to even "reasonably maximally exposed" individuals are acceptably small. Moreover, the total estimated risk for the whole population, considering all future exposures and assuming that the observed natural decline in PCB concentrations in the fish continues, is most likely zero cancers (at least 74% probability) due to contamination at the site, and the expected number is at most 1.5 cancers.

In quantitative health assessments such as this, we estimate risks to health as two broad categories: (i) excess risk of cancer, and (ii) risk of all adverse health effects other than cancer. Policy makers then decide what levels of risk of each type are acceptably small. On the federal

[1] Until the late 1970's, PCBs were used in a number of industrial and commercial products, such as carbonless copy paper. This paper was recycled at mills along the river, and residuals from this recycling have led to widespread, low-level contamination. PCBs have been detected in samples of the water, sediments, and fish of the Kalamazoo River and Portage Creek, and portions of the creek (below Cork Street, Kalamazoo) and river (from Portage Creek to Lake Allegan) comprise the Allied Paper, Inc./Portage Creek/Kalamazoo River Superfund Site. This risk assessment evaluates the area of the river extending from below Morrow Lake to the Allegan Lake dam. The area surrounding the river is generally rural although several cities and towns are located along its banks, in particular Comstock (population 11,800), Kalamazoo (population 80,000), Plainwell (population 4,000), Otsego (population 4,800), and Allegan (population 4,500) (MiIC, 2000).

[2] Since 1977, a fish consumption advisory has been in effect for this area. Nonetheless, anglers fish these waters, and some eat the fish they catch, despite "Do Not Eat" and "Limit Consumption" advisories. Our assessment, then, is an analysis of anglers' actual fish consumption rates, as extensively characterized by the State of Michigan in a comprehensive survey of Kalamazoo River anglers (MiCPHA, 2000 a,b,c).

level, policy makers at U.S. EPA consider "plausible high-end"[3] individual estimates of site-related excess lifetime cancer risk of up to 10 in 100,000 to be acceptably small. For this Kalamazoo site, our risk assessment finds that the plausible, high-end, estimate of individual cancer risk is 1.7 in 100,000, and so acceptably small.[4] For risks of all health effects other than cancer, we and other analysts calculate a "hazard index," and policy makers at U.S. EPA determine that site-related hazard indices smaller than 1.0 are acceptably small. The plausible high-end hazard index due to PCBs at this site is 0.81, again acceptably small, and indicative of no significant risk to health.

On the State level, for a reasonably maximally exposed individual (RMEI) cancer risk estimate, Michigan typically is ten-times more stringent than U.S. EPA, considering risks of up to 1 in 100,000 to be acceptably small. However, it is not clear whether this additional stringency is appropriate for sites, such as this one, in which the most probable cancer risk estimate for the entire affected population is zero cancers (and in which the RMEI estimate, at 1.7 in 100,000, is only slightly larger than 1 in 100,000) . Certainly, since the entire population of anglers herein studied, including those ingesting fish despite long-standing fish consumption advisories, are not at significant risk to health, the benefits of additional clean-up are expected to range from none to negligible.

This assessment has been produced to supplement and place in perspective the "Final Human Health Risk Assessment" (HHRA) for the Allied Paper, Inc./Portage Creek/Kalamazoo River Superfund Site (MiDEQ, 2000) produced by the Michigan Department of Environmental Quality (MiDEQ). That HHRA is, in most respects, a screening-level assessment. It evaluates hypothetical populations, uses maximum values of measured concentrations (either in soil or sediment at individual locations, or in individual fish), and uses parameter estimates (for example, for dermal absorption) specifically selected as upper bounds suitable for use in such screening-level assessments. In addition, the MiDEQ (2000) HHRA does not distinguish between different areas of the floodplain of the Kalamazoo River that are distinct in location, ownership, and potential uses. The MiDEQ (2000) HHRA also fails to mention a genuine,

---

[3] Plausible high-end estimates are those located at the 90th percentiles of the risk distributions herein constructed, taking full account of both variability and uncertainty in both exposures and toxicities, as explained below in the text. The choice of 90th percentile is in accord with EPA's Guidelines for Exposure Assessment (U.S. EPA, 1992b), and with EPA's draft guidance for probabilistic risk assessments (U.S. EPA, 1999).

[4] Because the doses of PCBS involved here, even for those consuming substantial amounts of Kalamazoo River fish, are more than 1,000 times smaller than doses shown to cause cancer in laboratory rodents, it is also plausible that no excess risk of cancer is presented by the site.

substantial, key uncertainty upon which its risk assessment rests. This uncertainty is the assumption that tiny levels of PCBs in fact cause cancer in people.[5]

While screening-level assessments can be useful, they have limitations. In particular, when the screening-level calculation suggests some scenario has risks that are unacceptably high, it is impossible to say whether that result arose because of a real-life high-risk situation, or because the screening-level calculations were overly conservative. This ambiguity was the case for the HHRA — it was impossible to say whether the source of the high estimates was the actual situation on the Kalamazoo, or instead the nature of the analysis applied. What is required in such situations is an analysis in greater depth that evaluates the scenarios with more realism and takes account of the uncertainties. Here we supplement the HHRA by performing some analyses in much more detail. So doing, and then comparing the results of our detailed assessment with those of the HHRA, we find that many of the results of the HHRA are unrepresentative, unreliable, or both.

The largest source of human exposure to PCBs from the Kalamazoo River comes from ingestion of fish. To more fully explore exposures, our quantitative risk assessment takes full account of all the available measurements on the exposed population, and incorporates the variability among members of that population and the uncertainties inherent in any measurements. Rather than evaluate hypothetical populations, as was done in the MiDEQ (2000) HHRA, we here examine the population of anglers who actually eat fish from the Kalamazoo River, a population that has been extensively characterized in a thorough survey expressly designed for that purpose (MiCPHA, 2000 a,b,c). Our calculations assume that the future fish-eating populations will be similar to that surveyed with respect to patterns of fish consumption.

The Kalamazoo River Angler Survey (MiCPHA, 2000 a,b,c) allows estimation of how often and for how long anglers have eaten fish from the Kalamazoo River. In addition, it provides information on what fish they eat, and how much of each fish. In this assessment, that information is linked with the measurements of concentrations of PCBs in the various fish. These quantities, and others, are combined in a Monte Carlo[6] risk assessment that fully accounts for the variation among individuals (variability) and the uncertainties of measurement (uncertainty). The results are calculated in the form of distributions of doses that show what fraction of the population will be exposed to what dose, and how uncertain are those dose estimates. For example, the distribution of doses may be represented by a graph (Figure S-1) that

---

[5] This assumption is not known to be true. For example, despite considerable epidemiologic study, PCBs are *not* among the 87 substances (such as arsenic, asbestos, and benzene), mixtures (such as alcoholic beverages, tobacco smoke, and coal-tars), and exposure circumstances (such as aluminum production, coal gasification, and iron and steel founding) known to cause cancer in humans (International Agency for Research on Cancer, IARC, 2001).

[6] Monte Carlo simulation is a statistical method for calculating quantities — such as dose and risk — repeatedly, using inputs randomly selected from the probability distributions for those inputs, to generate a full range of plausible values.

shows what fraction of the population of fish-eating anglers has a dose lower than any specified dose.

Our assessment is a full probabilistic analysis for fish-eaters and hypothetical high-end exposure scenarios (screening-level assessments) for other exposure circumstances. The lifetime average dose of PCBs for someone eating fish caught in the river is highly likely (92% probability, taking account of both uncertainty and variability) to be smaller than the Michigan Environmental Science Board's Health Protective Value (HPV) of 0.05 μg/kg-day. Averaging over the period when anglers are actually eating fish (which may range from a year to a lifetime), it is less likely (49%) that a fish-eater would have a dose rate lower than 0.05 μg/kg-day. However, when account is taken of the overly protective nature of the HPV for the less sensitive portion of the population (Fischer *et al.*, 1998) by accounting for the uncertainties and variabilities in toxicity estimates and the effect of exposure duration, we find that the plausible high-end hazard index for an individual fish-eater (that is, the 90th percentile of the full distribution) is 0.81 — that is, the dose for such a fish-eater would not cause any risk of adverse effects.

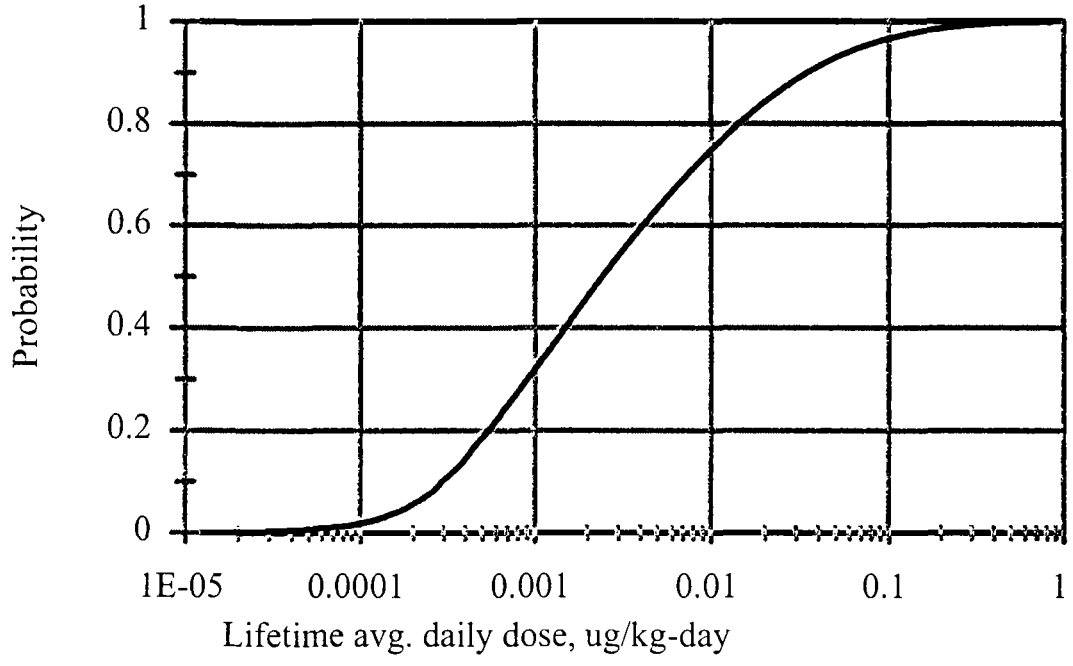

Figure S-1 Population variability in lifetime average daily dose, μg PCBs per kg body weight per day, with maximum likelihood estimates for uncertainty.

The range of cancer risk estimates for individuals is very wide, both because of the wide range of fish-eating habits and because of the inherent uncertainty in extrapolating results from studies in

laboratory animals to make predictions for humans. The major uncertainty is whether small quantities of PCBs indeed cause cancer in humans at all. Even assuming that PCBs do cause cancer at very low doses in humans at a rate that may be extrapolated from high dose experiments in laboratory animals, the median estimate for lifetime risk is below 1 in one million for a random fish-eater, and the 90th percentile is at 1.7 in 100,000 or less. This is a very small estimated risk, and may be compared with a background, lifetime risk of cancer from all causes of about 40% or 40,000 in 100,000. Moreover, with these conservative assumptions, the total effect in the whole population, adding up over all future times and assuming that the observed natural decline in PCB concentrations in the fish continues, is expected to be about 1.5 cancers, although the most probable number of cancers is zero (74% probability).

The calculations show that for the population of fish-eating anglers who start eating fish from contaminated areas of the Kalamazoo River in 1999, the best estimate of the *median* lifetime average dose rate of PCBs from eating those fish is 0.0025 µg/kg-day. This intake of PCBs for anglers eating Kalamazoo River fish is very small, both in absolute terms, and relative to the health-protective (that is, safe) value of 0.05 µg/kg-day. It is also substantially less than the 0.023 µg/kg-day from all sources that we estimate for current *non*-fish-eaters in the area, based on measured blood PCB concentrations. Further, the range of dose rates for different fish-eaters is fairly wide, because of the wide range of fish-eating habits. For example, the 10% of the population with smallest exposures will have a lifetime average dose rate below 0.00030 µg/kg-day, while fewer than 10% will have a lifetime average dose rate higher than 0.035 µg/kg-day.

These estimates of lifetime average dose rates are the best available; but we can also take account of the uncertainty in them. For the median estimate of 0.0025 µg/kg-day, for example, we can be 90 percent certain that the value lies somewhere between 0.0015 and 0.0053 µg/kg-day.

The calculations also allow evaluation of the dose rate during the period of exposure, as opposed to averaged over a lifetime. This period of exposure may range from a year to a lifetime in our calculations (because fish-eaters were assumed to eat fish for at least one year), so there can be a substantial difference between the two dose rates for any individual. For the dose rate during exposure, the best estimate of median dose rate is 0.048 µg/kg-day, while 10% of fish-eaters would have a dose rate below 0.012 µg/kg-day and 10% a dose rate above 0.27 µg/kg-day. While these estimates of average dose rates during exposure show some 49% (best estimate) of the population exceeding the health-protective value of 0.05 µg/kg-day, it must be borne in mind that the health-protective value itself is overly protective for the less sensitive portion of the population (Fischer *et al.*, 1998), such as men and women not of child-bearing age, and is for long-term exposure.

The total population of fish-eating anglers in the survey was less than 400. Adding in those people to whom anglers provided fish from the Kalamazoo, and those who were not surveyed, the best estimate for the total number of fish-eaters at any one time is likely to be about 6,870. Using the survey information applied to this larger population, about 1,004 people per year (best estimate) start eating fish from the Kalamazoo River, and about the same number per year

probably stop eating fish. Using these estimates, and the average dose estimates for the whole population of fish-eating anglers, we estimate that the annual impact of eating fish contaminated with PCBs from the Kalamazoo River amounts to between zero and approximately 0.038 cancers per year (median of the uncertainty distribution), with a 90 percent certainty of being less than 0.064 per year (assuming the EPA's upper-bound estimate for cancer potency of PCBs). This can be compared with the background rate of cancers (from all causes) in the same total population of those who ever eat fish from the Kalamazoo of about 400 cancers per year.

The concentration of PCBs in Kalamazoo River fish has been decreasing at a rate of about 5% per year. The risk estimates are all proportional to the concentrations in fish, so the estimated effects also decrease at about 5% per year. Taking account of this decrease, our best estimate of the *total* effect throughout the population of all fish-eating from the Kalamazoo over the entire future is less than 0.79 total cancers, spread over the entire future (and we can be 90 percent certain that the total effect is less than 1.7 cancers, again assuming the U.S. EPA upper-bound estimate for cancer potency of PCBs). Thus, even assuming the EPA's worst-case estimate for the potency of PCBs as human carcinogens, it is most likely that the total effect in the population is no extra cases of cancer (even though the mean estimate is 1.0 cancers). Such a result does not suggest a need for additional site clean-up.

The EPA's worst-case estimate for the potency of PCBs as human carcinogens does not really take account of all the uncertainties and variabilities involved. Those uncertainties and variabilities can be evaluated, conditional on an assumption that PCBs do in fact cause human cancer at low doses with a linear dose-response. When we evaluate these uncertainties and variabilities, and incorporate them into the calculations described, the median estimates for total population effect are smaller — about 0.0053 cancers/year and 0.11 total cancers ever. The upper 90th percentile estimates are larger, at 0.094 cancers/year and 2.2 total cancers occurring ever, and the expected total effect in the population is about 1.5 cancers. Even with these estimates, however, it is still more likely than not (there is an expected 74% probability) that there will be no cancers at all.

The MiDEQ (2000) HHRA shows, correctly, that exposures to PCBs over the majority of the floodplain of the Kalamazoo river, where soil concentrations average less than 2 mg/kg (ppm), do not pose unacceptable risks for residential exposure situations. The former impoundments at Plainwell, Otsego, and Trowbridge have higher concentrations of PCBs in the soils — which were formerly sediments — but these areas will not be used for residences, since they are state-owned wetlands. An analysis of the soil measurements shows that the upper 95th percent confidence estimates for the mean surface soil concentrations of PCBs within the impoundments are 36.0, 21.9, and 29.3 mg/kg for Plainwell, Otsego, and Trowbridge respectively. A screening-level risk assessment for the most highly exposed populations who come into contact with the soils in these three former impoundments — hunters and fishers — shows that their exposures are acceptably low. During the years that these hypothetical hunters or fishers are actually exposed, their dose rate of PCBs is unlikely to exceed 0.0024 µg PCBs/kg body weight per day (0.0024 µg/kg-day), well below the health-protective (that is, safe) value of 0.05 µg/kg-day.

Further, their lifetime risk of cancer from such doses would range from zero to no more than 2.8 in 1,000,000, well within acceptable limits, particularly for the small population sizes involved (probably fewer than 100 people).

Observations of the former impoundments have shown one instance of vegetable gardening by trespassers, possibly extending in that instance over a 20-year period. In a screening-level assessment, exposures of such vegetable gardeners are evaluated here to be potentially higher than for hunters and fishers — amounting to 0.15 μg/kg-day from vegetable ingestion, soil contact, and soil ingestion during the period they garden in the contaminated soils in the former impoundments, with the vegetable ingestion contributing the great majority. Their lifetime risk would range from zero to no more than 10 in 100,000 from these exposures.[7] The dose estimate exceeds the Michigan Health Protective value three-fold, and the risk estimate exceeds the Michigan standard for waste sites by a factor of about ten (although it is within U.S. EPA guidelines). However, the potentially exposed population is extremely small — probably fewer than 5 persons; and this exposure scenario is contingent on the failure of the State to enforce State Land Rules.

Overall, then, a detailed, probabilistic assessment of risks to health from PCBs at this site finds such risks to be very small. The plausible estimate for high-end (90^{th} percentile) cancer risk for an individual eater of Kalamazoo fish is 1.7 in 100,000; and the corresponding hazard index is 0.81. Moreover, even using conservative estimates, at most 1.5 cases of site-related cancer are expected in the entire population over the entire future, but the probability any cancers whatever is less than 26% . When public health risks, to individuals and to the entire population, are this small, action to further clean up a site is typically not warranted.

[7] A detailed, rather than screening-level, assessment would likely yield a lower risk estimate for a plausible high-end scenario.

1 Introduction

The Michigan Department of Environmental Quality has produced a "Final Human Health Risk Assessment" (HHRA) for the Allied Paper, Inc./Portage Creek/Kalamazoo River Superfund Site (MiDEQ, 2000). That risk assessment, however, is in most respects a "screening-level" assessment, in that it evaluates entirely hypothetical populations, uses maximum values of measured concentrations (either in soil or sediment at individual locations, or in individual fish), and uses parameter estimates (for example, dermal absorption) specifically selected as upper bounds suitable for use in such screening-level assessments.

Such screening-level assessments can provide assistance at some sites, particularly when they are adequate to rule out the presence of a problem. At this site, however, the use of a screening-level assessment is not helpful, particularly since the extent of the overestimates embodied in such an assessment cannot be evaluated. While the screening-level assessment shows risks that exceed typical regulatory maxima, it is not clear what groups, or how many, if any, individuals are actually exposed to a risk as high as the estimates obtained. Furthermore, the direct measurement of blood concentrations of PCBs in a sub-population that, according to the risk assessment, should be among the most highly exposed (that is, the sub-population that consumes Kalamazoo River fish), showed little effect from ingestion of fish (MiCPHA, 2000b), whereas the HHRA would predict a substantial effect.

To overcome some of the shortcomings of the screening-level HHRA, the document presented here performs a much more complete risk assessment for specific populations that are identified as those at highest risk — those exposed to fish from the Kalamazoo River and Portage Creek, and those potentially exposed to sediments in the former impoundments. We examine in detail the well-defined and well-studied population of anglers who regularly eat the fish they catch from the Kalamazoo (no anglers were found on Portage Creek, and hence no eaters of fish caught in Portage Creek). In addition, we examine (in somewhat less detail) the exposures to the sediments (now soil) of the former impoundment for those likely to be most highly exposed to them — hunters on the banks of the Kalamazoo, and trespassing gardeners using the highly fertile former impoundment soil to grow vegetables.

The detailed risk assessment for fish-eating anglers uses a probabilistic risk assessment (PRA), in contrast to the screening level approach of the HHRA. In both approaches, the same model is used for estimating what are the doses of PCBs and risks for an individual. The dose rate is calculated as the product of the concentration of PCBs in fish, the fraction of the PCBs that survive preparation and cooking, the amount of fish that is eaten by the individual per meal, and the number of fish meals per year. This continues for as long as the individual eats fish from the Kalamazoo, and the risk to that individual depends on both the dose rate during the period he eats

fish, the length of time he eats them, and the susceptibility of that person to the effects of PCBs. This description omits some details, but conveys the essence. However, choosing what values to use for these quantities is not so straightforward. We are somewhat uncertain about the concentration of PCBs in fish, because there are measurements in just so many fish. There is uncertainty about how much of the PCBs survive cooking. The amount of fish (and the types of fish) that any individual eats varies substantially from person to person. The number of fish meals per year eaten varies substantially from person to person. The length of time for which individuals eat fish from the Kalamazoo varies substantially from person to person.

In a screening level assessment, such as the HHRA, these uncertainties and variabilities are resolved (in a sense) by choosing values that correspond to a hypothetical relatively heavily exposed and susceptible eater of fish — for example, the uncertainty in concentration is resolved by choosing a high estimate; the variability in number of meals per year is resolved by choosing a number corresponding to a frequent eater; and the variability in length of time that an individual eats is resolved by choosing a period that corresponds to an upper end estimate of how long an individual might live in the area. The resultant estimate of dose and risk is thus a high-end estimate — but we have no idea how high end, or even if there could be any individual who would actually be so exposed. Moreover, this approach gives no clue as to the total effect on the whole population of fish-eaters.

In contrast, the PRA takes account of the full range of uncertainties and variabilities, and evaluates doses and risks for many individuals, using what is known as a Monte Carlo approach. For each individual, a value for each of the variables is chosen randomly from the range of values observed in the actual population, with a frequency corresponding to how often the value is observed. Thus the number of meals per year for each individual is chosen from the distribution of values for meals per year that was reported in an extensive survey of people eating fish from the Kalamazoo. Similarly the length of time an individual eats fish from the Kalamazoo was obtained from the distribution of times reported in that same survey. And so forth for the other variables. And correlations between variables are also taken into account.

This approach is repeated many times for many individuals (this is the "Monte Carlo" part), obtaining an estimate of dose and risk for each such individual. By doing this many times, we can build a picture of how often individuals will exceed a certain dose or risk — in other words, with each value of dose or risk we can associate an estimate of how likely it is for individuals to have that dose or risk. This is the desired result — we obtain estimates of dose and risk for a large number of individuals, and can ask what fraction of them have doses that exceed safe doses; or what fraction of them have risks exceeding acceptable values. In addition, because this approach evaluates many individuals, it allows estimates of the effect on the whole population of individuals involved, so that we can also estimate the total effect in the population while accurately accounting for all the differences between individuals.

There are myriad details involved in a Monte Carlo evaluation, including many essential ones that have not been included in the explanation just given. This document provides a complete

explanation of the risk assessment methodology, the input data (where they came from and how they are used), and provides a summary of the results. A complete reproduction of the assessment may be achieved with supplemental material provided along with this document in electronic form and listed in Appendix B and Appendix C. That material includes extensive tables of results (from which all the graphs and numerical results provided here were generated), and a complete set of the spreadsheet, program and support files.

2 *Exposure Scenarios, Pathways, and Receptors*

2.1 The physical setting

In 1971 the Michigan Department of Natural Resources (MDNR) identified PCB contamination in the water, sediments, and fish of the Kalamazoo River and Portage Creek, and in 1990 the U.S. EPA listed portions of the creek (below Cork Street, Kalamazoo) and river (from Portage Creek to Lake Allegan) on the National Priorities List (the Superfund list). The PCBs in the creek and river are present primarily due to the discharge of wastes from paper de-inking processes necessary for the recycling of carbonless copy paper which, until the mid-1970s, contained percent quantities of PCBs (BBL, 1992). This risk assessment is concerned with the area of the river extending from below Morrow Lake to the Allegan Lake dam. The area surrounding the river is generally rural although several cities and towns are located along its banks, in particular Comstock (population 11,800), Kalamazoo (population 80,000), Plainwell (population 4,000), Otsego (population 4,800), and Allegan (population 4,500) (MiIC, 2000). The land along the river is used for industrial, commercial, agricultural, residential, and recreational purposes (MiDEQ, 2000).

The Kalamazoo River and its environs have been extensively described in the Remedial Investigation (RI) and Feasibility Studies (FS) (BBL, 2000a,b) and associated memoranda. The exposure scenarios described here are the principal ones of concern identified in the Final Human Health Risk Assessment (MiDEQ, 2000) — exposure to the sediments of the former impoundments, and ingestion of contaminated fish from the Kalamazoo River.

2.2 The former impoundments

The Kalamazoo River was previously dammed at several locations within the study area for the production of hydroelectric power. Within the impoundments behind these dams, PCB-containing paper waste residuals settled into the river's sediments. When three of these dams (Plainwell, Otsego, and Trowbridge) were partially removed by the MDNR, the lower water levels exposed some of these sediments. Vegetation has since largely covered these former impoundment sediments, which form good substrates for plant growth. Paper waste residuals within the submerged sediments provide a continuing source of PCBs to the water and to the fish in the river. In areas of the former impoundments, additional PCBs are currently being introduced to the river by erosion of the banks, which are not very stable; and there is the potential for re-introduction during flooding. Although it is now possible for persons to come

into direct contact with the exposed sediments in limited areas, in general they are not easily accessible, and are thickly covered by vegetation.

2.2.1 The hunter/fisher scenario

The hunter/fisher scenario considers the case of a hunter or fisher who is exposed to the sediments in the former impoundment areas through dermal contact with, and ingestion of, the soil in these areas. Extended periods of hunting larger game animals (*e.g.* deer or bears) on these areas is highly unlikely, because of their small area and lack of such game animals. The most likely potential exposures are during relatively extended stays in blinds hunting seasonal wildfowl (principally during the fall hunting season). Several wood constructions having the appearance of abandoned wildfowl hunting blinds[8] were noted in the impoundment areas during a site visit in July, 2000. The total potentially exposed population is likely to be quite small, however, probably fewer than 100 persons.

Exposure routes to be considered are dermal exposure to surface soils, ingestion of soil, and inhalation of vapors from the soil (the last is shown to be negligible in Chapter 7). Dust emissions from the impoundment soils are negligible — the PCB-containing sediments are very fertile, resulting in rapid overgrowth by vegetation, so that wind stresses on the soil surface are negligible.

2.2.2 The trespassing gardener scenario

Observations revealed that a vegetable garden had been established in the former Otsego impoundment, an area with soil contaminated with PCBs. The contaminated sediment apparently makes fertile soil, so that there is an incentive to take over some of this public land for use as a private vegetable garden. Aerial photos indicated that a garden had been present at one location for a period at least 10 years, while contact with the gardener indicated that the period may have been as long as 20 years. On-site examination showed cultivation of vegetables and the presence of motorized gardening equipment. The gardener was required in August 2000 to abandon the garden on State property, and enjoined from harvesting any vegetables present; but the long delay before the State took notice indicates that a similar exposure could occur in future. While any such scenario thus depends on the failure of the State to enforce State Land Rules, it is evaluated because of the historical precedent for just such a failure.

[8] Such blinds are required by law to be removed at or before the end of the season, but these abandoned blinds were observed out of season.

2.3 *Fishing in the Kalamazoo*

From the 1940s through the 1960s the poor water quality of the Kalamazoo River made it generally unsuitable for fishing. Following the improvement in water quality that began in the 1970s and the stocking of the river by the MDNR, sustaining populations of a variety of game and rough fish have developed. The dams on the river provide both a good habitat for fish and ready access for anglers. The fishing habits and fish consumption patterns of anglers on the Kalamazoo River are well characterized in a survey conducted by the Michigan Department of Community Health (MiCPHA, 2000a). Between May 5, 1994 and September 30, 1994, survey teams encountered 1,090 anglers along the Kalamazoo River within Kalamazoo and Allegan counties and conducted face-to-face interviews with 938 of these individuals. Of the questions asked during the interviews, those directly relevant to assessing PCB exposures include: what type of fish caught from the creek and river are eaten by anglers and members of their households; how many meals of these fish are eaten annually; and how long each angler has been eating fish from the creek and river (although nobody was encountered fishing in the creek). The survey also included questions about the catching and consumption of snapping turtles from the creek and river. From the anglers questioned, 151 were later enrolled in a health survey and biological testing program, which included the measurement of PCB levels in blood serum samples.

Of the 938 anglers who responded to the survey questions, 345 indicated that they eat fish or turtles from the creek and river, and 294 said that a total of 807 other household members eat such fish or turtles (see spreadsheet Phase_1.wb3, Appendix B.12). Although the survey results include data concerning which fish types are consumed by the angler's household members, there are no data regarding the number of meals those household members eat. While it is likely that some anglers consume fewer fish meals than members of their households, it is also likely that the population of anglers themselves are generally the most highly exposed and so are exposed to the greatest amount of PCBs in fish from the creek and river. Household members consuming fish may include the children of adult anglers, but many such younger household members are unlikely to remain in the area permanently and continue angling from the Kalamazoo. Nevertheless, the survey results and hence this risk assessment, include relatively high proportions of anglers who must have started fishing at a very young age (less than 10) if the information they provided on duration of fishing and age is correct.

2.4 *Combinations of scenarios*

This document evaluates the scenarios discussed in Sections 2.2.1, 2.2.2, and 2.3 separately, as though the persons involved are distinct and cannot partake of any combination of activities. This matches the way the HHRA (MiDEQ, 2000) has evaluated scenarios, but is not necessarily completely accurate. It is possible that the gardener also hunts or fishes, that the hunter also gardens on the impoundments or fishes, or that the fisher hunts or gardens on the impoundments.

However, there is essentially nothing to be gained by attempting to evaluate combinations of these scenarios. For the first two, we have evaluated a conservative estimate for doses and risks, and find that (at the upper bound) the hunter is exposed to doses and risks that are around 50-fold less than the upper-bound estimates for the gardener. The screening-level estimates of doses and risks for the gardener are comparable with the 90^{th} percentile estimates for the random fisher, but the more extreme fisher estimates dominate those for the gardener. Thus the risks of the hunting scenario are substantially smaller than the uncertainty for the gardening scenario, and the risks for the gardening scenario are in turn are less than the uncertainties and variabilities for the fishing scenario. Moreover, the total populations involved are substantially different also — for the gardening scenario, probably fewer than 10 people, for the hunting scenario almost certainly fewer than 100 people, but for the fishing scenario around 70,300 ever-eaters of fish from the Kalamazoo (about 6,870 current fish-eaters).

Thus any attempt to combine the scenarios would not add to the information presented for each scenario separately here. For the hunting and gardening scenarios, the results would be no more accurate for the hypothetical upper-bound individuals evaluated. Combining either or both of those two scenarios with the fishing scenario would simply get lost in the uncertainty and variability involved, and so add no meaningful information.

3 *Acceptable levels of risk*

What target risk is appropriate for the clean up of this Site? Risk assessors, risk managers, regulators, and others have struggled for years to define and defend what is meant by acceptable risk. Although it is relatively easy to determine what size of risks are *de minimis* — one in a million, one in a billion, one in a trillion, and so on — what level is clearly not *de minimis* has, for a variety of reasons, been more difficult to determine. Some regulatory decisions from U.S. EPA reveal that individual (MEI, or maximally exposed individual) risks may be as high as 1×10^{-4} or even 10^{-3} and still be considered *de minimis*. For example, in its Final Rule on radon emissions from phosphogypsum stacks (57 Fed. Reg. 23305-6, June 3, 1992), U.S. EPA writes "[an] estimated maximum individual lifetime risk of fatal cancer . . . less than the benchmark of 1×10^{-4} is . . . presumptively acceptable." Further, in its Guidance on Remedial Actions for Superfund Sites with PCB Contamination (August, 1990, p. 28), U.S. EPA writes: "For Superfund sites, the risk remaining after remediation should generally fall within the range of 10^{-4} to 10^{-6} individual excess cancer risk." Other decisions by U.S. EPA reveal similar or more permissive definitions of acceptable risk, especially, as discussed below, when aggregate, population impacts have been factored in.

Definitions of acceptable risk are typically thought to encompass at least two notions — risk to the reasonably maximally exposed individual (RMEI), and risk to the affected population. The latter concept appears to have been ignored in the HHRA (MiDEQ, 2000). As recognized by former U.S. EPA Deputy Administrator F. Henry Habicht (Habicht, 1992) and others, assessing the former without assessing the latter is stopping short of providing important information. It is easy to envision two similar sorts of waste-sites — both with roughly the same estimated risk to an RMEI, but the first with perhaps 100 times the estimated risk to the affected population of the second (because of differences in size of the affected populations, uses of the sites, and so on). It should be clear that the second site is the safer one; and it is clear that Agency decisions of the past have (as they should) turned on estimates of population risks as well as on estimates of individual risks (see for example Travis *et al.*, 1987). Indeed, U.S. EPA has deemed acceptable individual risks (of cancer) of up to 3×10^{-3} in circumstances in which population risks (also called population impacts, aggregate risk, total incidence, and so on) are low (see Table 2 in Travis *et al.*, 1987). For example, U.S. EPA has decided not to reduce by regulation risks from zinc oxide plants (individual upper-bound risk estimate of 3×10^{-3}); from secondary lead smelters (individual upper-bound risk estimate of 3×10^{-3}); and from elemental phosphorus plants (individual upper-bound risk estimate of 1×10^{-3}) — in each case because the Agency performed a population risk analysis, and found that the aggregate risks to the affected populations were too small to require reduction. The existing environmental risks from the industrial processes were thus deemed acceptable, even though, on an individual basis, they were as high as 3×10^{-3}.

For the Kalamazoo site, the potentially affected population is relatively small compared with the U.S. population. Even given current site conditions, aggregate risks to the population are near zero, and would be smaller still following any additional clean up (either naturally occurring or with human intervention). In such a situation, U.S. EPA policy allows individual risks of 10^{-4} or higher, as shown above. Moreover, the State of Michigan allows waste-site risks for maximally exposed individuals of up to 10^{-5}.

A useful scheme for consistent regulation of carcinogenic risks is that proposed by Kocher and Hoffman (1987). These analysts suggest three guideposts:

- Risks in the range of 10^{-1} to 10^{-3} are clearly significant, *de manifestis*, and need to be reduced by regulation regardless of cost:
- Risks in the range of 10^{-4} to 10^{-6} (or lower) are clearly insignificant, *de minimis*, and not deserving of reduction by regulation; and
- Risks that lie between these two ranges (between clearly significant and clearly insignificant) should be reduced following the principle of "as low as reasonably achievable" (ALARA).

4 *The central uncertainty in risk assessments for PCBs — toxicity values*

4.1 The fundamental problem

There is a genuine, substantial, key uncertainty upon which the entire risk assessment for the Kalamazoo River rests. That uncertainty, of course, is the assumption that tiny levels of PCBs in fact harm human health. Nowhere in the HHRA chapter on Uncertainty Assessment was there any mention of this key uncertainty. The HHRA assumes that very low levels of PCBs can affect human health in three ways In particular, it assumes that very small amounts of PCBs: (1) present a risk of cancer, (2) harm the developing nervous system, and (3) harm the immune system. It must be said that *none* of these assumptions is known to be true.

With respect to risk of cancer, for example, it should be noted that PCBs, even at high levels of exposure, are *not* known to cause cancer in humans — they are only *presumed* to do so based on extremely high-dose, lifetime feedings studies in laboratory rodents. As such, PCBs are one of the 684 substances or mixtures that have so far been shown to cause cancer in at least one species of laboratory animal (Gold and Zeiger, 1997). Moreover, of these 684 substances, fewer than 10% are known to be human carcinogens, and PCBs are *not* among these (International Agency for Research on Cancer, IARC 2001). U.S. EPA recognizes that PCBs are not known human carcinogens, referring to the human data as "inadequate" (IRIS, 2001). The Agency for Toxic Substances and Disease Registry (ATSDR, 2000) concludes, "Overall, the human studies provide some evidence that PCBs are carcinogenic."[1]

Importantly, it is not for lack of epidemiologic study that scientists conclude that PCBs are not established human carcinogens: some dozen epidemiologic investigations have examined cancer incidence or mortality in groups of people with PCB exposure and have found no strong or consistent evidence of an effect (*e.g.*, Bahn *et al.*, 1976; Bertazzi *et al.*, 1987; Brown, 1987;

[1] As a matter of policy, U.S. EPA and IARC regard PCBs are "probable human carcinogens." This designation is based on (i) the established carcinogenicity of PCBs when administered to laboratory rodents at high doses throughout their lifetimes, combined with (ii) the assumption that very low doses (such as those received by people affected by this site) would also be carcinogenic to laboratory rodents (though such doses have never been tested in bioassays), and (iii) the assumption that such doses would also be carcinogenic to humans. Since even massive doses of PCBs are not demonstrably carcinogenic to humans, many in the scientific community question to validity of these central assumptions.

Sinks *et al.*, 1992; Yassi *et al.*, 1994; Gustavsson and Hogstedt, 1997; Loomis *et al.*, 1997; Kimbrough *et al.*, 1999). The people studied were generally workers in electrical equipment manufacture or maintenance, in which industries very high exposures to PCBs could and did occur. Indeed, some of these investigations gathered data on workers' blood levels of PCBs, and found levels many-fold higher, even *hundreds of times* higher, than body burdens measured in the general population or in anglers in particular. Women in the general population have been studied repeatedly with regard to PCBs and breast cancer risk (Wolff *et al.*, 1993; Krieger *et al.*, 1994; Hunter *et al.*, 1997; Moysich *et al.*, 1998; Helzlsouer *et al.*, 1999; Dorgan *et al.*, 1999), but results on the whole do not suggest an effect: results of two of the more comprehensive studies (Hunter *et al.*, 1997 and Helzlsouer *et al.*, 1999) suggest that higher body burdens of PCBs are associated with *lower* than expected risks of breast cancer.

Only in mega-dose studies in laboratory rodents, then, have PCBs been demonstrated to reliably and reproducibly induce cancer. Moreover, there are striking differences between the magnitudes of the doses studied in the laboratory and the doses received by people eating contaminated foods. The *lowest* daily dose of PCBs received by rats in the bioassay used by U.S. EPA to estimate PCBs' cancer potency is equivalent to a human dose of 0.35 milligrams of PCBs per kilogram of body weight (mg/kg) (IRIS, 2001); but a person eating an average of 15 grams of fish per day containing 2 ppm PCBs, for example, would receive an average daily dose of only 0.00043 mg/kg — some *800 times smaller*. Given the non-mutagenic nature of PCBs, it is highly questionable whether even laboratory rodents would be at risk of cancer from doses this tiny. The assumption that *people* would be at such risk is more tenuous still. Even for *established* human non-mutagenic carcinogens, such as alcohol, no responsible scientist would predict that, because 3 drinks per day for life increases a person's risk of cancer, the equivalent of 0.00375 drinks per day (3 ÷ 800) also presents a significant, actionable risk of cancer.

In workers exposed to high levels of PCBs and other chlorinated compounds, certain non-cancer health effects such as chloracne and other dermatologic signs sometimes developed (although less so as the hazard was recognized and controlled in the 1940's), but few other effects have been reliably linked to PCBs. With regard to environmental exposures to PCBs, there is little evidence of harm to adults.[2] Among women eating contaminated Lake Ontario fish (containing contaminants such as mercury and some pesticides, besides PCBs), there is some suggestion of shortened menstrual cycles, but no evidence of interference with conception or risk of spontaneous miscarriage (Mendola *et al.*, 1995, 1997; Buck *et al.*, 1997). Regarding the shortened menstrual cycles, a finding that awaits confirmation, the authors note that "they are not likely to be clinically relevant" (Mendola *et al.*, 1997).

---

[2] The epidemics of contaminated cooking oil-related illnesses in Japan and Taiwan, referred to as Yusho and Yu-Cheng, respectively, are interpreted by most scientists as indicative primarily of the toxic effects of polychlorinated dibenzofurans (PCDFs), rather than of PCBs. U.S. EPA does not rely on these epidemics for derivation of cancer potencies or reference doses for PCBs (IRIS, 2001).

There is mixed evidence of neurobehavioral deficits, decreased birth weight, and growth deficits in children exposed to PCBs and other contaminants *in utero*. Whereas one group of investigators has reported apparently persistent effects in Michigan children (Fein *et al.*, 1984; Jacobson *et al.*, 1990a, 1990b, 1996), investigators studying children in North Carolina saw only transitory effects (Rogan and Gladen, 1991; Gladen and Rogan, 1991), and still other investigators have found no adverse effects in infants born to women who consumed fish from Lake Michigan (Dar *et al.*, 1992) or from Lake Ontario (Mendola et al., 1995). Of course, in all of these studies, fish were also variably contaminated by mercury and other compounds, so that the studies could not evaluate only PCBs. Some studies from Europe seem to confirm findings of decreased weight at birth, decreased growth rate in the first few months after birth, decreased performance on tests of cognitive abilities, or transient deficits in neurologic measures (Lanting *et al.*, 1998; Patandin *et al.*, 1998; Patandin *et al.*, 1999). The studies also support earlier observations that consumption of breast milk by infants does *not* produce adverse effects (but, rather, beneficial effects) despite the amounts of PCBs that may be thereby consumed. Performance and growth deficits are typically slight, and within the range of the general population. The significance of these early-life changes at adolescence or adulthood has not yet been assessed. Whether observed deficits are due entirely or largely to PCBs, as opposed to other chemicals that typically accompany dietary PCBs, is not yet certain, but experimental work with nonhuman primates supports a PCB effect (Barsotti and van Miller, 1984; Schantz *et al.*, 1989; Schantz *et al.*, 1991). Further, recent work by Stewart *et al.* (2000) finds that newborns' performance, as assessed by the Neonatal Behavioral Assessment Scale, seemed to be impaired in relation to levels of highly chlorinated PCBs in cord blood, but not in relation to cord blood levels of other contaminants, such as lead, DDE, hexachlorobenzene, mirex, or lightly chlorinated PCBs, nor in relation to maternal hair levels of mercury. However, recent work by Longnecker *et al.* (2000), finds that PCBs do *not* disrupt thyroid function (which disruption can harm the developing nervous system) in human fetuses exposed to PCBs, even as prenatal exposure of rats to PCBs does affect their thyroid function postnatally. The investigators speculate that rodents may be more susceptible than humans in this regard.

The mixed evidence from these various studies on *in utero* effects is difficult to interpret. In its *Toxicological Profile for PCBs*, ATSDR (2000) concluded, "No firm conclusions can be made regarding growth and development of children and environmental exposures to PCBs, although perinatal exposure to high concentrations of PCBs and structurally related chemicals, as occurred in *Yusho* and *Yu-Cheng*, affects birth weight and growth during early life." This state of knowledge, or lack thereof, should give responsible toxicologists pause. Whether any reliable risk assessment for PCBs can yet be performed with a developmental-neurologic endpoint in mind is highly debatable. Further, with respect to children (as opposed to fetuses) or adults exposed to PCBs *via* contaminated soil, sediments, or fish, no studies suggest that neurological or neurobehavioral effects might occur.

Finally, we infer that the HHRA (MiDEQ, 2000) relies in part on the results of experiments in which rhesus monkeys were dosed with moderate levels of PCBs for two years or more. Two types of effect were observed: dermatologic reactions and alterations in immunologic parameters.

The implication that dermatologic reactions might occur in people exposed to low levels of PCBs does not comport with the evidence from occupational epidemiology, where dermatologic effects usually did not occur at even moderate (and sometimes high) exposures. The changes in immune system parameters observed in monkeys after two years of exposure to PCBs were statistically significant, but did not appear to have any clinical (*i.e.*, functional) significance. Although studies are scarce, no signs of damage to the immune system have been reported in cohorts exposed occupationally to PCBs (Emmett *et al.*, 1988a, 1988b), and neither immunologic damage nor chloracne or related conditions has been observed in populations with environmental exposure to PCBs.

Overall, then, the abundant scientific literature describing the human experience with PCBs does not establish that clinically significant harm is caused by environmental contamination of the degree existing at and near the Kalamazoo River site. Responsible risk assessments should make this plain.

4.2 A partial solution for cancer potency

Given the central uncertainty about whether PCBs cause cancer in humans, it is not possible to provide unconditional estimates of the probability for the exposures at this site to cause cancer. However, it is possible to provide upper bound estimates *conditional on PCBs causing cancer in humans and conditional on the linear-low-dose hypothesis.* The first of these is obviously conservative, and the second is also considered to be conservative even in the stronger form in which the linearity is extended all the way from zero dose up to the lower end of the range of animal bioassay experimental doses.

Upper-bound point estimates for the carcinogenic potency of PCB mixtures have been provided by Cogliano (1996) using the standard type of approach of the U.S. EPA — these are the estimates reported in IRIS (2001). Such estimates are conditional on the PCBs causing cancer in humans, and are also conditional on the strong form of the linear-low-dose hypothesis. The highest upper-bound value estimated for the potency of environmental mixtures of PCBs is 2 kg-day/mg. This value is used in the non-probabilistic parts of this document, in analyses that are by nature highly conservative in estimating doses, and also used to some extent to interpret the probabilistic dose estimates obtained for the fish ingestion scenario.

However, the estimates of Cogliano (1996) provide no indication of the conservatism of the analysis used, and do not capture the full uncertainties and variabilities even within the adopted hypotheses. Therefore, we have analyzed the available bioassay data and applied a methodology that does take into account these uncertainties and variabilities. These probabilistic estimates for carcinogenic potency have been incorporated in the analysis of the fish ingestion route, to illustrate more completely the range of variability and uncertainty.

4.2.1 Analysis of laboratory animal bioassays

The carcinogenic potency of PCBs was evaluated in a non-probabilistic way for U.S. EPA by Cogliano (1996) to obtain upper bound estimates for carcinogenic potency for various mixtures of PCBs. We use the same bioassay data, with a similar but more comprehensive ED_{10} approach for analyzing the animal data, but use the approach of Crouch (1996) to fully incorporate experimental uncertainties and the uncertainties of interspecies extrapolation.[3]

There are fifteen available lifetime bioassays in laboratory rodents that are considered suitable for evaluation of dose-response curves. These bioassays were of commercial PCBs, either Aroclors or Clophens, and their results are incorporated in five publications. Kimbrough *et al.* (1975) fed 100 ppm Aroclor 1260 in the diet to female Sherman rats. NCI (1978) fed 25 ppm, 50 ppm, and 100 ppm Aroclor 1254 to groups of male and female Fischer rats. Schaeffer *et al.* (1984) fed 100 ppm Clophen A-30 and A-60 to groups of male Wistar rats. Norback and Weltman (1985) fed Aroclor 1260 to male and female Sprague Dawley rats at 100 ppm initially, decreased to 50 ppm after 16 months, then to zero after 24 months. Mayes *et al.* (1998) fed commercial PCBs 1260, 1254, 1242, and 1016 to male and female Sprague Dawley rats at 0, 25, 50, and 100 ppm in the diet.[4]

We evaluate the liver tumor responses from these bioassays using the re-evaluation of the liver slides reported in Moore *et al.* (1994) for the first four publications (seven bioassays), just as in Cogliano (1996). Results from Mayes *et al.* (1998) are used unchanged, since their liver evaluations were equivalent to those of Moore *et al.* (1994). The results we used are those given in Cogliano (1996), where animals dying before the first tumor appeared are censored (this information is not available in Mayes *et al.*, 1998). The information used on all fifteen bioassays is provided in Table 4.1.

These data were fitted by dose-response curves of the form

$$p(d) = 1 - \exp\left(-\left(a_0 + a_1 d + a_2 d^2\right)\left(t/t_0\right)^3\right)$$

where the terms are:

---

[3] Calculations are performed in the spreadsheet PCB_cancer_dose_response.wb3, Appendix B.5.

[4] Cogliano (1996) obtains liver tumor information on this study from Brunner *et al.* (1996), and dose information from Keenan and Stickney (1996). We used the published mean dose information from Brunner *et al.* (1998, Table 2), which has minor differences from Cogliano (1996).

$p(d)$ probability for a liver tumor at dose rate d,
t effective period on study,
t_0 standard period on study for lifetime experiments (104 weeks),
a_0, a_1, a_2 parameters ≥ 0.

The dose-response curve was limited to the given exponential linear-quadratic form, even for experiments with three dosed groups, to agree with the recommendation of the peer review panel cited by Cogliano (1996). For experiments with just one dosed group the quadratic term was omitted in order to obtain unique ED_{10} estimates.

The technique for fitting was maximum likelihood, assuming that the observed results are binomial observations with a probability $p(d)$, as is the standard U.S. EPA approach (Anderson *et al.*, 1983). As in that standard approach, if the dose-response curve fitted inadequately, the highest dose group was recursively omitted until the fit was adequate.[5] Our approach to fitting the dose-response curves differs from that of Cogliano (1996) in two ways:

a. All curves were fitted simultaneously, so that a control group was used only once when it was the control for multiple experiments (as in Mayes *et al.*, 1998, for example).
b. The time factor $(t/t_0)^3$ was incorporated to reduce all experiments to a standard lifetime. This is a standard practice (Anderson *et al.*, 1983) that was not incorporated by Cogliano (1996).

From the parameter estimates a_0, a_1, and a_2 for each experiment, an estimate of ED_{10} (the dose at which the tumor incidence is increased by 10%) may be obtained as:

$$ED_{10} = \frac{-2\ln\left(1-0.1\exp\left(a_0\right)\right)}{a_1 + \sqrt{a_1^2 - 4\ln\left(1-0.1\exp\left(a_0\right)\right)}}$$

and hence the potency (in the experimental rat strain) as $0.1/ED_{10}$. The potencies calculated in this way have relatively small uncertainties (that is, there is generally less than a factor of 2 difference between the maximum likelihood estimate and an upper confidence bound,[6] see Table 3-1 of Cogliano, 1996). Moreover, since these potency estimates depend on linear extrapolations from the ED_{10}, which lies close to or within the experimental range of non-zero doses, using different dose-response curves to interpolate the experimental results should provide very similar estimates.

---

[5] Cogliano (1996), followed this approach, as indicated by the results documented in Appendix A of that reference.

[6] The lower confidence bound could be zero, so that the uncertainty in this direction is, in a sense, infinite. See Section 6.9 for further discussion.

There are multiple experiments for each Aroclor, and each Aroclor has a different potency in each rat strain. The different potencies in the different experiments for each Aroclor may be due to differences in the congener profile of the PCBs tested in each experiment but labeled as the same Aroclor, or to differences between the rat strains tested, or to a combination of these. For example, it is known that there are differences between various Aroclors that are assigned the same name. The Aroclor 1254 tested by Mayes *et al.* (1998) and representative of perhaps 0.5% of the total PCB market is known to have a tetrachloro-dibenzo-(p)-dioxin (TCDD) toxicity equivalent (TEQ) concentration approximately double that of a sample of Aroclor 1254 representative of about 15% of the commercial market of Aroclor 1254, so it is possible that "the response for females in this study overrepresents that would have been expected for ordinary Aroclor 1254" (Mayes *et al.*, 1998). Again, although Mayes *et al.* (1998) and Norback and Weltman (1985) both tested Aroclor 1260 in Sprague Dawley rats, there was a clear difference in response for the females. The weaning body weights for the rats were also a factor of two different between these experiments.

To obtain an overall estimate of potency suitable for extrapolation to humans, all the experiments were analyzed simultaneously for the potency of Aroclor 1254, and for the potencies of Aroclors 1260, 1242, and 1016 relative to Aroclor 1254. For this analysis, it was assumed that these relative potencies would remain similar within experiments performed by the same experimenters with the same rat strain, but that the absolute potency of Aroclor 1254 could vary between experiments. This approach assigns all the differences between tests to variations in the sensitivity of rat strains, rather than differences in Aroclor congener profiles, and thus may be overestimating the variability between strains but underestimating uncertainty (see also Section 6.9).

Table 4.1 Liver tumor response to lifetime dosing of PCB mixtures.						
	Mixture	Dose ppm	Dose [a] mg/kg-d	Effective period [b]	Animals with liver tumor	Number of animals
Mayes *et al.* (1998)						
Female Sprague Dawley rats	Control	0	0	104	1	85
	1260	25	1.4	104	10	49
		50	2.8	104	11	45
		100	5.8	104	24	50
	1254	25	1.4	104	19	45
		50	2.9	104	28	49
		100 [c]	6.1	104	28	49
	1242	50	2.8	104	11	49
		100	5.7	104	15	45
	1016	50	2.7	104	1	48
		100	5.4	104	6	45
		200	11.2	104	5	50
Male Sprague Dawley rats	Control	0	0	104	7	98
	1260	25	1	104	3	50
		50	2	104	6	49
		100	4.1	104	10	49
	1254	25	1	104	4	48
		50	2	104	4	49
		100	4.3	104	6	47
	1242	50	2	104	1	50
		100	4	104	4	46
	1016	50	2	104	2	48
		100	4	104	2	50
		200	8	104	4	49

Table 4.1 Liver tumor response to lifetime dosing of PCB mixtures.						
Kimbrough *et al.* (1975)						
Female Sherman rats	Control	0	0	100	1	187
	1260	100	5.37	100	138	189
NCI (1978)						
Male Fischer rats	Control	0	0	104	0	24
	1254	25	1.25	104	1	24
		50	2.5	104	1	24
		100	5	104	3	23
Female Fischer rats	Control	0	0	104	0	23
	1254	25	1.25	104	1	24
		50	2.5	104	2	24
		100	5	104	1	24
Schaeffer *et al.* (1984)						
Male Wistar rats	Control	0	0	115	8	120
	A-30 [d]	100	5	115	16	128
	A-60 [d]	100	5	115	114	125
Norback and Weltman (1985)						
Male Sprague Dawley rats	Control	0	0	126	0	31
	1260	100	5	126	5	40
Female Sprague Dawley rats	Control	0	0	126	1	45
	1260	100	5	126	41	46

[a] Average dose rates. For Mayes *et al.* (1998, Table 2, mean values), as published. For Kimbrough *et al.* (1975), averaged over the published dose curve. Otherwise as estimated in Cogliano (1996) — (ppm in diet) × 0.05 — corresponding to an average food consumption of 5% of body weight per day. For Norback and Weltman (1985) the initial dose levels (given in the table) were reduced after 16 months. However, we use the initial dose level, as did Cogliano (1996).

[b] Study duration, as given by Moore *et al.* (1994) for bioassays other than Mayes *et al.* (1998).

[c] As in Cogliano (1996, Table A-2), this dose group was omitted from the analysis, since the fit of the dose-response formula was inadequate with it included.

[d] Schaeffer *et al.* (1984) provide the fractions of the tested Clophen's that are biphenyl, monochlorobiphenyl, dichlorobiphenyl, and so on. A-30 is similar to typical Aroclor 1016, and A-60 to Aroclor 1260. They are here treated as equivalent to those Aroclors.

4.2.2 Potency estimates

The potency estimates for the fifteen bioassays were parametrized in terms of potency estimates β for Aroclor 1254, and the ratios of potencies $R_{60/54}$, $R_{42/54}$, and $R_{16/54}$, where $R_{60/54}$ is the ratio of the potency of Aroclor 1260 to that of Aroclor 1254 in the same sex and strain, and similarly for the other ratios, as follows:

Mayes et al. (1998),

Female Sprague Dawley rats		Male Sprague Dawley rats	
1260	$\beta_1 \cdot R_{60/54}$	1260	$\beta_2 \cdot R_{60/54}$
1254	β_1	1254	β_2
1242	$\beta_1 . R_{42/54}$	1242	$\beta_2 . R_{42/54}$
1016	$\beta_1 . R_{16/54}$	1016	$\beta_2 . R_{16/54}$

Kimbrough et al. (1975)

Female Sherman rats	
1260	$\beta_3 \cdot R_{60/54}$

NCI (1978)

Male Fischer rats		Female Fischer rats	
1254	β_4	1254	β_5

Schaeffer *et al.* (1984)

Male Wistar rats		
A-30	$\beta_6 \cdot R_{16/54}$	(Treating A-30 as equivalent to Aroclor 1016)
A-60	$\beta_6 \cdot R_{60/54}$	(Treating A-60 as equivalent to Aroclor 1260)

Norback and Weltman (1985)

Male Sprague Dawley rats		Female Sprague Dawley rats	
1260	$\beta_7 \cdot R_{60/54}$	1260	$\beta_8 \cdot R_{60/54}$

This approach yields the three ratios $R_{60/54}$, $R_{42/54}$, and $R_{16/54}$ and eight estimates of the absolute potency of Aroclor 1254. The maximum likelihood estimates for all these parameters were obtained simultaneously (see the spreadsheet PCB_cancer_dose_response.wb3, and the additional information in Appendix B.5) and are:

$R_{60/54}$	0.455	β_1	0.242	β_4	0.0245	β_7	0.0313
$R_{42/54}$	0.269	β_2	0.0230	β_5	0.019	β_8	0.509
$R_{16/54}$	0.020	β_3	0.567	β_6	0.679		

where all potencies are expressed in kg-day/mg. The lower potency of Aroclor 1260 relative to 1254 is also supported by the smaller effects seen for Aroclor 1260 in a sub-lifetime experiment in Sherman rats (Kimbrough *et al.*, 1972). The variation in potency estimates for Aroclor 1254 between different experiments (essentially, between different rat strains and sexes) gives an idea about the range of variation to be expected between individuals or distinguishable population groups, and so was used to represent variability in an animal (or human) population. It was

assumed to be adequately represented by a lognormal distribution with median 0.106 kg-day/mg and natural logarithmic standard deviation 1.62 (a factor of 5.06), the values obtained from these eight estimates.[7]

4.2.3 Extrapolation of potencies to humans

For interspecies extrapolation from rats to humans, the best available estimate is that of Crouch (1996). This indicates that, conditional on a chemical being a carcinogen in both species, such an interspecies extrapolation may be represented by a lognormal uncertainty distribution with median unity (when potencies are expressed on a kg-day/mg scale), and natural logarithmic standard deviation of 2.2 (a factor of 9.1).

Thus, conditional on PCBs being human carcinogens at all, and conditional on a linear dose-response relationship below the ED_{10}, the potency of Aroclor 1254 in humans is best estimated from the rat data by a value of 0.106 kg-day/mg, with a lognormal uncertainty distribution with median unity and natural logarithmic standard deviation 2.2, together with a lognormal variability distribution with median unity and natural logarithmic standard deviation 1.62. For a randomly chosen individual, the U.S. EPA upper-bound point estimate of 2 kg-day/mg is at the 85.9^{th} percentile of the total uncertainty distribution (the variability distribution becomes an uncertainty distribution for a randomly chosen individual). For the population average potency (averaging over the variability distribution — the population average potency is 0.393 kg-day/mg at the median of the uncertainty distribution), the U.S. EPA upper-bound point estimate of 2 kg-day/mg is at the 77^{th} percentile of the uncertainty distribution. For an individual, the U.S. EPA upper-bound point estimate of 2 kg-day/mg is at the 96.5^{th} percentile of the variability distribution (at the median of the uncertainty distribution). Finally, the expected value of the distribution, averaging over both variability and uncertainty gives an estimate of 4.4 kg-day/mg, more than twice as high as the U.S. EPA upper-bound point estimate — so that using our approach on a population that was all identically dosed would give expected value estimates of cancer more than twice as large as EPA's "upper bound."

---

[7] The uncertainties in the values for the individual estimates of ED_{10} are negligibly small compared with the variation between values — the individual variability — that is incorporated here. We here mean the uncertainty as indicated by the difference between lower confidence limits and maximum likelihood values for ED_{10}; the upper confidence limits can be infinity (see Section 6.9). This procedure of representing the variability by a lognormal distribution is a refinement of the typical U.S. EPA approach of taking a geometric mean of values found in different experiments.

4.2.4 Application of potencies for individual Aroclors

The potency estimates derived here were used in the probabilistic analysis for the fish ingestion pathway. The average Aroclor composition of the PCBs measured in each type of fish is shown in Section 6.2.7. Dose estimates due to ingestion of each type of fish were obtained separately, and the potencies of the individual Aroclors applied. For Aroclor 1254, the potency estimate just described was used. For Aroclors 1260, 1242 and 1016, the ratios $R_{60/54}$, $R_{42/54}$, and $R_{16/54}$ were assumed to apply also to individual humans, independent of the variabilities and uncertainties — *i.e.* the potencies for the individual Aroclors were assumed to be always in the same ratio for individuals. For Aroclor 1248, the ratio $R_{48/54}$ was assumed to be 0.635 (half way between Aroclor 1242 and 1254). While these values are subject to some uncertainty, it is likely to be small compared with the uncertainties already documented, and so is ignored here (the uncertainty range is estimated in Section 6.9..

4.3 A partial solution for non-cancer effects

As has been discussed, there is limited evidence of harm to health from environmental (as opposed to occupational) levels of exposure to PCBs. Various attempts have been made to estimate a dose rate for PCBs that can be considered safe for long-term exposure. For the non-probabilistic part of this document, we adopt the estimate of 0.05 μg/kg-day endorsed by the Michigan Environmental Science Board (MESB) (Fischer *et al.*, 1998), and apply it to the dose rate occurring during the period of exposure. This value is also used for comparison purposes in examining the doses estimated in the probabilistic analysis of the fish ingestion route. However, the MESB (Fischer *et al.*, 1998) noted:

> In September 1995 the MESB concluded that the HPV of 0.05 μg/kg/day proposed in the 1993 draft Protocol for a Uniform Great Lakes Sport Fish Advisory (GLSFATF, 1993) was sufficiently protective of the most susceptible portion of the population (Fischer *et al.*, 1995). Michigan's proposal to adopt that advisory approach for 1998 for women of childbearing age and children represents a cautious approach and has the support of the present MESB Panel. The 1995 MESB report also indicated that the same HPV appeared overly protective for the less sensitive portion of the population and that less restrictive advice could be applied in recognition of the benefits derived from consumption of fish in moderate quantities. The view of the current Panel is that there are no new data that require an alteration of this conclusion and there is merit in continuing to give less restrictive advice to consumers of sport-caught fish in the less vulnerable portion of the population.

The value of 0.05 μg/kg-day is thus considered to be over-conservative for the less sensitive part of the population, such as men. To take account of the variation of sensitivity in the population, and the uncertainty in extrapolating from animal laboratory experiments to the human

population, the probabilistic part of this document also makes fuller use of the available data, as described in the following sections. This probabilistic analysis is based on the same data as used by the U.S. EPA in deriving a Reference Dose (RfD) (IRIS, 2001), and the ATSDR in deriving a Minimal Risk Level (MRL) (ATSDR, 2000).

4.3.1 Probabilistic evaluation of minimum effect levels

The Reference Dose (RfD) estimated by U.S. EPA is intended to be an estimate (with uncertainty spanning perhaps an order of magnitude) of a daily exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime (Barnes and Dourson, 1988). As such, the RfD cannot provide numerical estimates of risk for non-cancer health effects, but only describes the exposure conditions that are not likely to cause adverse effects (Gaylor and Kodell, 2000). For the probabilistic analyses performed here, we would ideally like to estimate the probability for deleterious effects at various lifetime average dose rates of PCB mixtures. Gaylor and Kodell (2000) provide a summary of various investigations that provide estimates for such probabilistic analysis that approaches this ideal, and indicate how the factors considered can be combined.

The RfD for Aroclor 1254 is based on experiments in rhesus monkeys fed 0, 5, 20, 40 and 80 µg/kg-day for a total of about 6.5 years (Tryphonas *et al.*, 1989, 1991a, 1991b; Arnold *et al.* 1989, 1993a, 1993b). The LOAEL was chosen to be 5 µg/kg-day based on ocular exudate, prominence and inflammation of the Meibomian glands, and distortion in nail bed formation; decreased antibody (IgG and IgM) response to sheep erythrocytes was also observed (IRIS, 2001). No dose-response information for the ocular exudate, prominence and inflammation of the Meibomian glands, or distortion in nail bed formation has ever been published (private communications between D.L. Arnold and E.A.C. Crouch, October 2000 through January 2001), although some measures of these effects were significantly changed (Arnold *et al.* 1993a). Since the publication of the RfD by the U.S. EPA, the same group of investigators have reported on reproductive and infant effects (Arnold *et al.* 1995), endometriosis (Arnold *et al.* 1996), and post-reproductive and pathological findings (Arnold *et al.* 1997) in the same group of monkeys, and on infant feeding studies in other groups of monkeys (Arnold *et al.* 1999). The major finding of these subsequent studies was the marked effect on reproduction — a reduced rate of impregnation and an increase in fetal mortality with increasing dose (Arnold *et al.* 1995). The infants of the dosed monkeys experienced the same clinical signs as their dams, and immunological findings that paralleled the results for the dams. The infants' clinical signs were less severe than those of their dams and generally appeared after weaning (Arnold *et al.*, 1995).

Dose-response curves for the effect of PCB dosing on IgG and IgM responses to sheep erythrocytes were published (Tryphonas *et al.*, 1989, 1991b), although whether those effects are adverse is not clear. In their evaluation of a Minimum Risk Level (MRL), which is based on these immunologic responses, ATSDR (2000) states that "Interpretation of the adversity of this effect is complicated by the lack of data on immunocompetence and the essentially inconclusive findings in the other tested end points; however, support for the 0.005 mg/kg/day LOAEL is

provided by mild clinical manifestations of toxicity at the same dose." A recent review considers that "In particular, there is a need for good correlative data between chemically induced changes in immune function measurements and changes in host resistance to specific disease. Until such correlations are established, interpretation of the observed shifts in lymphocytes and their subsets is only speculative." (Tryphonas, 1998).

To obtain the RfD for Aroclor 1254, U.S. EPA divided the assumed LOAEL of 5 μg/kg-day by four uncertainty factors. The factors are meant to account for:

Sensitive individuals — that is, intra-species variability from one individual to another,
Interspecies extrapolation — the uncertainty of extrapolating between species,
LOAEL to NOAEL extrapolation — to estimate a dose that has no effect, based on the lowest dose that has been observed to have an effect, and
Subchronic to chronic extrapolation — to account for effective exposure periods longer than those used to determine the LOAEL.

We examine these factors, and summarize the distributions suitable for use in their place for a probabilistic assessment.

4.3.2 Sensitive Individuals — intra-species variability

A nominal factor of 10 was used by U.S. EPA to account for sensitive individuals. This factor is incorporated to account for human variability, and a similar factor of 10 was used by ATSDR (2000) in deriving the MRL. The factor has been evaluated by Dourson and Stara (1983), and Gaylor and Kodell (2000) summarize that evaluation by a lognormal distribution with median unity and a natural logarithmic standard deviation 1.64. This distribution is a variability distribution, and is used as such. The 10-fold factor then is at the 92nd percentile of the distribution.

4.3.3 Interspecies extrapolation and its uncertainty

A factor of 3 was applied by U.S. EPA to account for extrapolation from rhesus monkeys to humans, and a similar factor was used by ATSDR (2000) in deriving the MRL. The default factor here is 10, but it was considered that "A full 10-fold factor for interspecies extrapolation is not considered necessary because of similarities in toxic responses and metabolism of PCBs between monkeys and humans and the general physiologic similarity between these species" (IRIS, 2001, Aroclor 1254). U.S. EPA also appears to believe that "In general, Rhesus monkeys have shown adverse effects to PCB mixtures at doses 10-fold lower than in other species" (IRIS, 2001, Aroclor 1248). There is no published empirical distribution for the monkey-human extrapolation; all that is available is the general interspecies uncertainty distribution that Gaylor and Kodell (2000) interpret as lognormal with median unity and natural logarithmic standard

deviation of 1.66, based on Calabrese and Baldwin (1995). On this scale, the 3-fold factor applied is at the 74.6th percentile, and the usual factor of 10 is at the 91.7th percentile. We adopt the published general interspecies uncertainty factor.

4.3.4 LOAEL to NOAEL extrapolation and its uncertainty

A factor of 3 was applied by U.S. EPA to account for extrapolation from a LOAEL to a NOAEL. The factor 3, rather than the usual 10-fold factor, was applied because "the changes in the periocular tissues and nail bed see [sic] at the 0.05 mg/kg-day are not considered to be of marked severity" (IRIS, 2001, Aroclor 1254). This statement, however, confuses severity (in the sense of how adverse is the effect) with the dose-response for the occurrence of that effect (to what degree does the effect occur, or in what fraction of animals, or some combination). It is notable that the ATSDR (2000) used the full factor of 10 for the LOAEL to NOAEL extrapolation in deriving a NOAEL. For the probabilistic analysis, we use the LOAEL/NOAEL statistics compiled by Pieters *et al.* (1998) for chronic studies, represented as a lognormal uncertainty distribution with median 4.3 and logarithmic standard deviation of 0.53. This is slightly different from the choice of Gaylor and Kodell (2000), who based their estimate of a median 3.5 with logarithmic standard deviation 0.60 on Abdel-Rahman and Kadry (1995), whose database was much smaller (24 chemicals versus 175). The EPA-applied factor of 3 is only at the 25th percentile of our chosen distribution, while the usual factor of 10 is at about the 94.4th percentile.

The use of this LOAEL/NOAEL extrapolation ensures that our estimates are ultimately for a NOAEL — a dose at which no adverse effect is expected. Thus the approach we adopt can be used and interpreted in the same sense that an RfD is usually used and interpreted.

4.3.5 Subchronic to chronic extrapolation

U.S. EPA applied a factor of 3 to account for extrapolation from a subchronic to chronic study. It was noted that the study (which extended over about 6.5 years) was for about 25% of the lifespan of rhesus monkeys, so that a factor of 3 was used in place of 10. However, application of this factor appears problematic. It is reported that during the three year pre-breeding phase of the experiment, 90% of the monkeys "had attained a satisfactory qualitative pharmacokinetic steady state regarding the concentration of PCB in their adipose tissue" by 25 months (Arnold, 1993a), and a six year study would generally be considered chronic in almost any species. ATSDR (2000) also applies no factor to account for extrapolation to longer periods in deriving the MRL. The best available evaluation of the subchronic to chronic uncertainty distribution considered all studies exceeding 1 to 2 years to be chronic (Pieters *et al.*, 1998). For the probabilistic analysis, therefore, no additional uncertainty factor was applied for a sub-chronic to chronic exposure period.

However, in the probabilistic analysis we have to account for exposures of differing durations. It appears likely that the mechanism of action of PCBs for non-cancer effects depends on the concentration of PCBs in human tissues, or equivalently on the body burden of PCBs. While the rhesus monkeys dosed with Aroclor 1254 reached approximately equilibrium body burdens within about 25 months, it appears that human metabolism is considerably slower, so that humans dosed at constant dose rates would not reach equilibrium body burdens within a lifetime. Thus the NOAEL in humans will decrease with the length of exposure, since at a constant dose rate the maximum body burden will increase with increasing exposure duration. The accumulation of PCBs with age in humans is described in more detail in Sections 6.3.2 and 6.11.2. Figure 6.22 shows the increase of body burden with age, assuming constant dose rate from birth. This figure may also be interpreted as the increase in body burden with time since first exposure, assuming a constant dose rate.

For the probabilistic risk assessment therefore, the NOAEL derived by extrapolation from the rhesus monkey experiment LOAEL will be assumed to correspond to an exposure of 25% of a standard lifetime in humans, corresponding to the 25% of a lifetime in the rhesus monkey experiment. To estimate a NOAEL for other periods of exposure, we use the inverse of the curve shown in Figure 6.22, normalized to unity at an exposure period of 17.5 years (25% of the nominal human lifetime of 70 years). The result is shown in Figure 4.1 (see spreadsheet PCB_congener_data.wb3, Appendix B.13). The NOAEL estimated by extrapolation from the rhesus monkey experiment corresponds to 17.5 years exposure duration. For shorter durations, the NOAEL corresponding to the same maximum body burden is higher, rising to 5.8 times higher for 1 year exposure. For longer durations, the NOAEL is lower, down to 0.64 of the 17.5 year NOAEL at an exposure duration of 70 years.

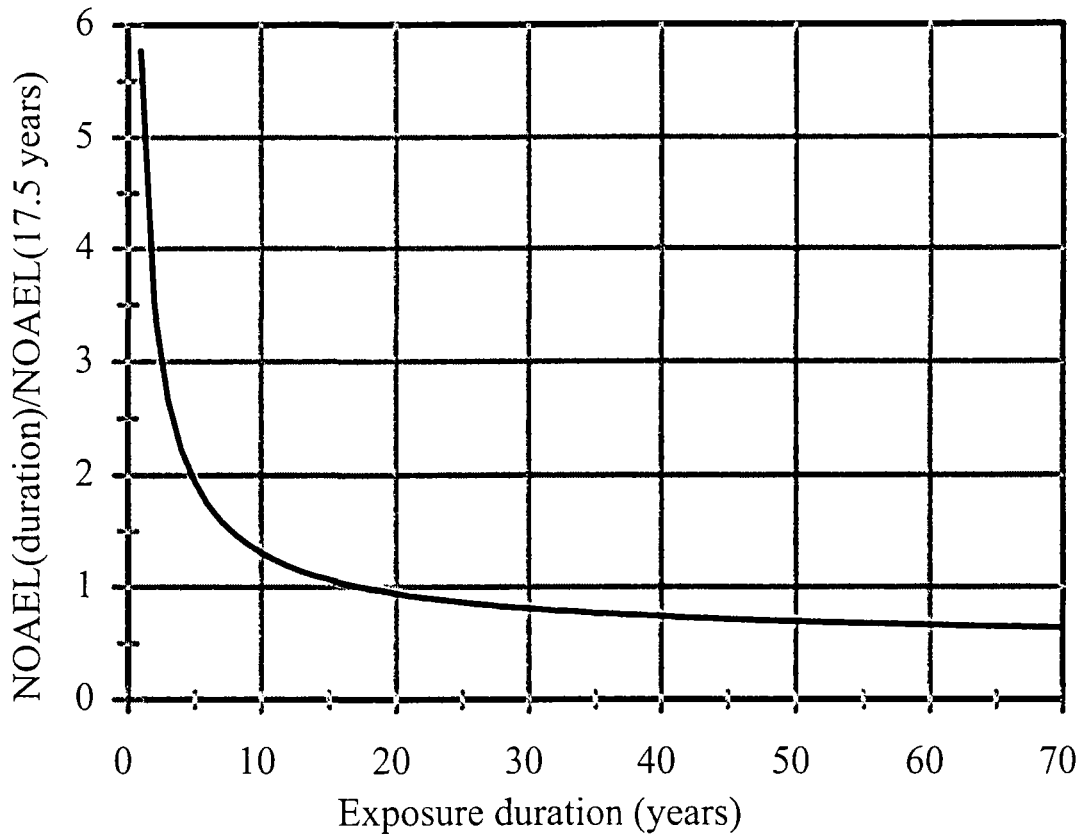

Figure 4.1 NOAEL for differing exposure durations, relative to the NOAEL for 17.5 years.

In summary, the probabilistic analysis uses three lognormal distributions, two of which are uncertainty distributions and one a variability distribution, together with a curve to adjust for differing exposure durations. The variability distribution has median unity and logarithmic standard deviation 1.64. The uncertainty distributions have medians unity and logarithmic standard deviation 1.66 (interspecies), and median 4.3 with logarithmic standard deviation 0.53 (LOAEL to NOAEL). Combining the uncertainty distributions gives a lognormal with median 4.3 and logarithmic standard deviation 1.743. Combining all distributions for a randomly chosen individual with exposure duration of 17.5 years would give a lognormal distribution with median 4.3 and logarithmic standard deviation 2.393. The 300-fold factor obtained by U.S. EPA by using the four factors above (or the ATSDR using three factors) is at the 96.2th percentile of this uncertainty distribution for the randomly chosen individual. In addition, the exposure duration is taken into account using Figure 4.1. The difference between 17.5 year and 70 year (lifetime)

exposure is a factor 1.56. Incorporating this factor, the 300-fold factor obtained by the U.S. EPA is at the 94.4th percentile of the uncertainty distribution for a randomly chosen individual.

To high accuracy (better than 0.1% for all durations greater than 5 years), the calculated curve may be represented by:

$$\frac{\text{NOAEL(duration } t)}{\text{NOAEL(17.5 years)}} = \frac{1}{h_m\left(1-\exp\left(-h_0\left(1-\exp\left(-h_1 t\right)\right)-h_2\left(1-\exp\left(-h_3 t\right)\right)-h_4 t\right)\right)} \qquad (4.3)$$

where the terms, their values, and their dimensions,[8] are:

t		exposure duration (T),
h_m =	1.808546	(dimensionless),
h_0 =	0.067816	(dimensionless),
h_1 =	0.964302	per year (T^{-1}),
h_2 =	0.517510	(dimensionless),
h_3 =	0.077284	per year (T^{-1}),
h_4 =	0.020221	per year (T^{-1}),

The above discussion strictly applies to Aroclor 1254, which by several measures appears to be the most toxic of the commercial PCB mixtures. An RfD for Aroclor 1016 was also derived (IRIS, 2001), based on a different end point and with slightly different modifying factors. That RfD was three-fold higher than the RfD for Aroclor 1254. No RfDs have been published for Aroclors 1242, 1248, 1260, or for environmental mixtures of PCBs, in the absence of suitable experimental results. For Aroclor 1248, the best available study had a death in the lowest dose group, and was judged unsuitable as the basis for an RfD. However, the LOAELs in other experiments have all been higher than the 5 μg/kg-day LOAEL adopted for Aroclor 1254.

At this site, the PCBs are a mixture of congeners that are approximated by a mixture of Aroclors. For the primary exposure pathway of interest, through ingestion of fish, the dominant component of the best-approximating Aroclor mixture is Aroclor 1254. We shall use the probability distribution derived above as an approximation, probably slightly conservative, for the total PCBs at this site. The curve defined by Figure 4.1 has been derived for a mixture of Aroclors

[8] Following standard scientific practice, terms in equations in this document represent physical quantities, not just numbers, so that there is never any need for conversion factors within equations. Physical quantities have a magnitude and dimensionality that are expressed by a numerical value and a unit; the equations, however, do not impose requirements for any particular system of units. Dimensionality is represented by combinations of the seven standard dimensions — mass, length, time, electrical current, temperature, amount of substance, and luminous intensity. Only the first three, represented by M, L, and T, are needed in this document.

corresponding to a 75%:25% mix of the Aroclor distributions measured in bass and carp respectively, and is insensitive to the variations in Aroclor distribution found in fish at the site.

4.4 Do children form an especially sensitive subgroup?

As ATSDR (2000) states:

> Children are exposed to PCBs in the same manner as the general population, primarily via consumption of contaminated foods, particularly meat, fish, and poultry (Gunderson 1988). Exposure also may occur by transfer of PCBs that have accumulated in women's bodies to the fetus across the placenta. Because PCBs are lipophilic substances, they can accumulate in breast milk and be transferred to nursing infants. Transfer across the placenta, although it may be limited in absolute amounts, is of great concern because of the effects of PCBs on sensitive immature tissues, organs, and systems, with potentially serious long-lasting consequences. Transfer of PCBs via breast milk can be considerable and, like prenatal exposure, has the potential to contribute to altered development.
>
> Although embryos, fetuses, and nursing infants may be exposed to relatively high amounts of PCBs during sensitive periods of development, it is not known if the susceptibility of children to the health effects of PCBs differs from that of adults. Younger children may be particularly vulnerable to PCBs because, compared to adults, they are growing more rapidly and generally have lower and distinct profiles of biotransforamtion [sic] enzymes, as well as much smaller fat depots for sequestering the lipophilic PCBs.

Children are not considered as a special subgroup in any of the analyses performed in this document. Such treatment might be appropriate if children were especially susceptible to the effects of PCBs (so that the health protective value or minimum effect levels used in the analyses were set too high to protect them), or if there was some sub-group of children particularly heavily exposed to the PCBs from the site that requires separate analysis. Neither situation applies here, according to the best available scientific evidence.

The ATSDR (2000) statement, and that of the MESB (Fischer *et al.*, 1998, see Section 4.3), summarize the scientific evidence on susceptibility and on the protectiveness of the health protective value. The most direct evidence comes from the studies of Arnold *et al.* (1995) on reproduction and infant findings in female monkeys dosed at 5, 20, 40, and 80 µg/kg-day Aroclor 1254 for 6.5 years, and their offspring. The offspring were exposed *in utero* and through their dams' milk prior to weaning. Although they were exposed during weaning to total PCB intakes (on a body weight basis) that were substantially higher than their dams, the infants experienced toxicological effects that were generally similar to their dams, except that their "clinical signs were *less* severe than those of their dams and generally appeared after weaning" (Arnold *et al.*,

1995, emphasis added). In a more recent study in which infant rhesus and cynomolgus monkeys were fed 7.5 μg/kg-day PCBs of a congener composition designed to mimic that in human breast milk for their first 20 weeks of life (pre-weaning), there were few differences between treated and control groups as a consequence of the PCB ingestion, and those were judged to be transient, minor or biologically insignificant (Arnold *et al.* 1999). A surprising finding of the last study was the differences in the blood and lipid PCB concentrations that were observed between animals administered the PCBs in their liquid diet or in corn oil. The mean blood concentration in the animals administered PCBs in corn oil was at least 6-fold higher than in the animals administered PCBs in the liquid diet. This finding is currently unexplained, but suggests marked, vehicle-dependent differences in absorption of PCBs across the gastrointestinal tract.

While infants may be relatively highly exposed to PCBs as a consequence of breast-feeding, the monkey experiments indicate that the health protective value adopted is protective of these infants also, so that infants do not need to be considered as a special group. There also appears to be no particular childhood group that is at risk for substantially increased ingestion rates of PCBs from this site. Children are less likely to be exposed to the former impoundments than adults, and if exposed are likely to be exposed less often. While Michigan encourages adults to take their children hunting, it is unlikely that children will spend as much time hunting as adults. Thus the hunting scenario is limited to evaluation of adults, who are the highest exposed group. Similarly, children are unlikely to spend nearly as much time tending vegetable gardens as adults.

In the fishing population, there is direct (albeit limited) evidence of the lack of special status for childhood exposure in results from Phase II of the Kalamazoo River Angler Survey (MiCPHA, 2000a,c). In that survey, the blood PCB concentration of twelve children under the age of 15 were measured. Seven of the children (ages 11.6, 12.6, 13.5, 14.2, 14.2, 14.7, and 14.7) had not eaten fish from the Kalamazoo river, and had blood total PCB concentrations of 0.5, 1, ND (<0.1), 0.5, 0.5, 0.9, and 1 ppb respectively. Five children (ages 9.5, 10.9, 12.9, 14.7, and 14.8) had been eating such fish for a claimed 9, 10, 5, 1 and 4 years respectively. Their blood total PCB concentrations were 0.6, ND (<0.1), 1.6, 0.5, and 1.1 ppb respectively, indicating no substantial early childhood increment in blood PCB concentrations, and by implication no particularly high childhood exposures.

5 *The former impoundments*

5.1 Current surface soil exposure-point concentrations

Exposure-point concentrations[1] for the hunter-fisher scenario are estimated from the transect sampling measurements of surface soils from the former impoundments.[2] Any individual hunter/fisher is likely to be exposed to relatively large areas of the former impoundments over the long run. Moreover, individual measurements on these transects are here regarded as repeated samples of similar materials (the paper waste materials mixed with river sediment) distributed over the former impoundment area, rather than as samples of distinct areas with different characteristics. Both points of view lead to the statistical treatment of these data as repeated samples of an exposure concentration that varies from one exposure occasion to another, so that the average exposure concentration may be estimated by averaging the transect samples taken within the former impoundment areas. Since the three different former impoundment areas are physically distinct, at different distances down river, they are treated as separate areas.

Except possibly for the trespassing gardener (see Section 5.3), exposures to the soils in the former impoundments would be to the surface soil only, so only the surface soil measurements are used here. The laboratory reported concentrations for each of Aroclors 1221, 1232, 1016, 1242, 1248, 1254, and 1260 (listed here in order of molecular weight, and so also of average

[1] In this chapter, when we use the term "concentration" for the PCB content of soil, we strictly mean the "mass fraction" of the soil. The two terms are generally used interchangeably in this way, although the concepts are different — a concentration is a mass per unit volume, whereas the mass fraction is a mass per unit mass. We nevertheless continue the generally accepted usage, except in the definition of terms in equations, where it is important to be precise.

[2] All data and calculations are in the spreadsheet Impoundment_data.wb3, Appendix B.7.

chlorination level; ATSDR, 2000).[3] There were no detects of any of Aroclors 1221 and 1232 in any of the transect samples, so the concentration of those two Aroclors has been taken to be zero. The lack of detection of these Aroclors, and its treatment as their absence, are consistent with the more rapid environmental removal of the lower-chlorinated Aroclors. The laboratory also reported a summary statistic consisting of the sum of the detected Aroclor concentrations if any were detected, and a value of ½ the lowest detection limit of any Aroclor if no Aroclors were detected. That summary statistic is used for convenience in the discussion of sample selection in the following four paragraphs, but the subsequent statistical analysis for the exposure point concentration probably takes better account for the detection limits of all the Aroclors ever detected in the transect samples.

All the surface soil sample locations obtained in the transect sampling (duplicates counted as a single location), 42 for Plainwell, 41 for Otsego, and 76 for Trowbridge, were examined, and sample locations outside the former impoundments were omitted, since this analysis is for the former impoundment soils alone; inclusion of concentration measurements from outside the impoundment would dilute the exposure point concentration estimate.

At Plainwell, the impoundment is reported to have been at elevation 712 ft, but there is a clear demarcation in the samples somewhere between 712.25 ft and 713.98 ft. No sample above 713 ft is reported to contain the grey clay-like material, and most samples below that elevation are so reported. Of the 12 surface sample locations above this elevation, 10 are non-detect for PCBs, with two containing 0.30 and 0.86 mg/kg. Elevation 713 ft was taken as the demarcation line for the former impoundment.

At Otsego, of the 14 surface soil sample locations above elevation 683 feet, the reported impoundment elevation, 11 were non-detect, one contained trace levels (0.048 mg/kg), and two had elevated concentrations of 2.3 and 9.3 mg/kg. The 4 sub-surface measurements at the same locations showed non-detects. None of these samples was described as containing the characteristic grey clay-like deposits, whereas almost all the samples taken below elevation 683 ft were so described. Elevation 683 ft was taken as the demarcation line for the former impoundment.

At Trowbridge, 11 of the 12 sample locations above elevation 669 ft, the reported impoundment level, were non-detect, with the last having a trace concentration of 0.098 mg/kg. Once again, no

[3] The identification of aged environmental samples of PCBs as mixtures of Aroclors is necessarily problematic, since the congener mix in the samples may not correspond to any such Aroclor mixture. The laboratory used a constrained least-squares algorithm to obtain a best estimate for an approximately equivalent mixture of the Aroclor standards used, but the exact specification of effective detection limits for those Aroclors imputed to be absent from the sample is not straightforward. The procedures we adopt are designed to take reasonable account of the uncertainties involved. See also the discussion of unquantified uncertainties in Section 6.9.

sample above this elevation was described as containing grey clay deposits, and most samples below this elevation contained such deposits. Elevation 669 ft was taken as the demarcation line for the former impoundment.

To better account for the detection limits on all the ever-detected Aroclors (1016, 1242, 1248, 1254, 1260), the possible range of total PCB concentrations was obtained for each measurement by summing the measured concentrations of Aroclors 1016, 1242, 1248, 1254, and 1260, first assuming that a non-detect corresponded to zero concentration, and second assuming that non-detects corresponded to a concentration equal to the detection limits for each Aroclor. The result is an upper and lower bound on the total PCB concentration from each sample location. Duplicate samples at the same location were combined: the average was taken if both were detects; otherwise if either was a detect, that value was used; otherwise the lowest of the two detection limits was used as the detection limit for the combined sample. This procedure gave upper and lower bounds[4] on the total PCB concentration for 30 sample locations at Plainwell, 27 at Otsego, and 64 at Trowbridge.

These data were then used to estimate an upper 95th percent confidence limit on the mean total PCB concentration of surface soils within each impoundment. Calculating such a confidence limit requires some evaluation of the underlying statistical distributions of the values of concentrations at the individual sampled locations. The typical approach is to evaluate whether the distribution of values is lognormal, and if it is, to use the procedure of Land (1971, 1973, 1974, 1975, 1988; Lyon & Land, 1999) to obtain an estimate of the 95th percent confidence limit on the mean. If the distribution of values is not lognormal, any estimate obtained by Land's procedure is likely to be highly biased. In that case, the upper 95th percent confidence limit on the mean is typically estimated as though the distribution were normal, by using a t-statistic based estimator. The latter approach is asymptotically unbiased for almost all underlying distributions, and is extremely robust against differing underlying distributions. This typical approach is limited in requiring point estimates for all measurements that were non-detects, and the standard estimate is to use ½ the detection limit.

As a first test, the means of the upper and lower bounds were computed for each sample location (this corresponds to using ½ the detection limit for all non-detect samples), and the distribution of this mean value was plotted and tested to see if it is consistent with a normal or lognormal distribution. The distribution of these concentrations within the former Plainwell impoundment is consistent with being lognormal, as evident on a lognormal probability plot (Figure 5.1), with a Shapiro-Wilk statistic (Royston, 1982, 1993, 1995) of p_{sw}=0.46. The arithmetic mean and sample standard deviation of the 30 sample values obtained in this way are 17.6 mg/kg and 22.2 mg/kg respectively.

[4] Since these concentration estimates are themselves used in the subsequent distributional analysis, the upper and lower "bounds" are used more as an indicator of the range of values that have high likelihood than as absolute limits.

58 Charles Street Cambridge, Massachusetts 02141
617-225-0810 FAX: 617-225-0813 www.CambridgeEnvironmental.com

For Otsego and Trowbridge, a single lognormal distribution is clearly not an adequate fit when using ½ the detection limit, as evidenced by Shapiro-Wilk statistics $p_{sw} = 0.006$ for Otsego, and $p_{sw} = 1.6 \times 10^{-5}$ for Trowbridge using the point sample estimates (½ detection limits), and by the non-linearity clearly apparent on probability plots (Figures 5.2 and 5.3).

For a sample of point values from a lognormal distribution, various methods are available to estimate the upper 95th confidence limit (UCL95) on the mean. For example, applying the jackknife approach using a minimum variance unbiased estimator (Singh *et al.*, 1997) to the averages of upper and lower bounds for each sample for the Plainwell samples gives an estimate of 27.9 mg/kg; the standard likelihood approach gives an estimate of 36.8 mg/kg, and the procedure of Land (1971, 1973, 1974, 1975, 1988; Lyon & Land, 1999) gives 39.9 mg/kg.

The statistics generated by the typical approach to estimating exposure point concentrations are given in Table 5.1.

Table 5.1 Statistics for the typical approach to estimating exposure point concentrations — former impoundments.

	Number of samples	Total PCB conc. (ppm) [a]			p-value for log-normal [c]	UCL95 estimate [b]	
		Mean	SD	Max		Normal	Lognorm
Otsego	27	13.7	12.5	40.4	0.006	**17.8**	60.9
Plainwell	30	17.6	22.2	102	0.46	24.5	**39.9**
Trowbridge	64	16.7	19.0	91.3	1.6×10^{-5}	**20.7**	81.4

[a] Using ½ detection limit for all non-detected Aroclors that were ever detected.
[b] Upper 95th percent confidence limit on the mean, assuming the underlying distribution is normal (Normal) or lognormal (Lognorm). The former uses the t-statistic, the latter the procedure of Land (1971, 1973, 1974, 1975, 1988; Lyon & Land, 1999). **Bold** figures indicate the estimate that should be selected using this typical approach.
[c] Shapiro-Wilk statistic (Royston, 1982, 1993, 1995) for the logarithms of the sample values.

The rightmost column can only be used for Plainwell, where the distribution is consistent with being lognormal. For Otsego and Trowbridge, where the distributions are clearly non-lognormal, the correct application of the typical approach would require the use of the estimates obtained in the second column from the right.

To further evaluate the exposure point concentrations at this site, we made an effort to identify the distributions of values, and to take better account of the non-detect values. The bias was to expect a lognormal distribution, or a combination of lognormal distributions, since

environmental and contamination data are often distributed lognormally or as a sum of lognormals.

For a statistical sample of ranges of values, as available here, or for a sample from a non-normal or non-lognormal distribution, the only standard statistical approach available for estimating confidence limits on the mean is the likelihood method. Land's method (1971, 1973, 1974, 1975, 1988; Lyon & Land, 1999) can only be applied to point estimates of values from a lognormal distribution, omitting the uncertainty induced by the finite detection limits. While a jackknife method could be applied, it would first require some estimation procedure on which to apply it — and the minimum variance unbiased estimation approach is not available for non-point estimates or for mixtures of distributions. The only available procedure that can be used is thus the likelihood approach, as used below.

For the Plainwell samples (incorporating the range of values for each sample), a single lognormal distribution provides an adequate fit (as judged by the Shapiro-Wilk statistic applied to the averages, as described above, and by visual examination of a probability plot, Figure 5.1). Using the likelihood method applied to the ranges of values, the mean concentration is estimated to be 19.0 mg/kg, with a 95th upper confidence limit on this mean of 36.0 mg/kg.

Examination of the logarithmically transformed data shows that the Otsego dataset may be adequately represented by the sum of two lognormal distributions, while the Trowbridge dataset requires three lognormals for an adequate fit (although one degenerates to a constant). Figures 5.1, 5.2 and 5.3 show the distribution models fitted to the three sets of data using maximum likelihood methods, with the uncertainties due to the non-detect values incorporated on the plot (so that each measurement is represented on the plot by a line ranging from its minimum likely total PCB content to its maximum likely content).[5]

The distribution models used here are sums of one, two, or three lognormals. The general case (a sum of k lognormals) has a cumulative probability distribution for PCB mass fraction m in soil of:

$$F(m) = \sum_{i=1}^{k} f_i \Phi\left(\left(\ln(m) - \mu_i\right)/\sigma_i\right) \quad \text{with} \quad \sum_{i=1}^{k} f_i = 1 \quad \text{and} \quad 0 \le f_i \le 1 \tag{5.1}$$

[5] Figures 5.1, 5.2, and 5.3 plot the logarithm of the mass fraction measured in the samples against a transform of their rank order. The transform is chosen so that a lognormal distribution shows as a straight line on this plot — it is the inverse normal of $(i-3/8)/(n+1/4)$ for the concentration measurement with rank i of n total samples. This value is a close approximation for the expected location of this i^{th} rank measurement if the distribution is lognormal (Cunnane, 1978). Completely non-detect samples have lower bound estimates extending to infinity on the left.

where the f_i, μ_i, and σ_i are parameters of the distribution, and Φ is the cumulative normal function. For a set of samples with lower and upper bounds (l_j, u_j) on PCB mass fraction, for $j = 1$ to n, the loglikelihood function is

$$L = \sum_{j=1}^{n} \ln\left(\sum_{i=1}^{k} f_i \left(\Phi\left(\left(\ln(u_j) - \mu_i \right) / \sigma_i \right) - \Phi\left(\left(\ln(l_j) - \mu_i \right) / \sigma_i \right) \right) \right) \tag{5.2}$$

The mean of the distribution is

$$M = \sum_{i=1}^{k} f_i \exp\left(\mu_i + \sigma_i^2 / 2 \right) \tag{5.3}$$

The loglikelihood is evaluated, and the maximum likelihood estimates and confidence limits on the mean M obtained, in the spreadsheet Impoundment_data.wb3 (Appendix B.7).

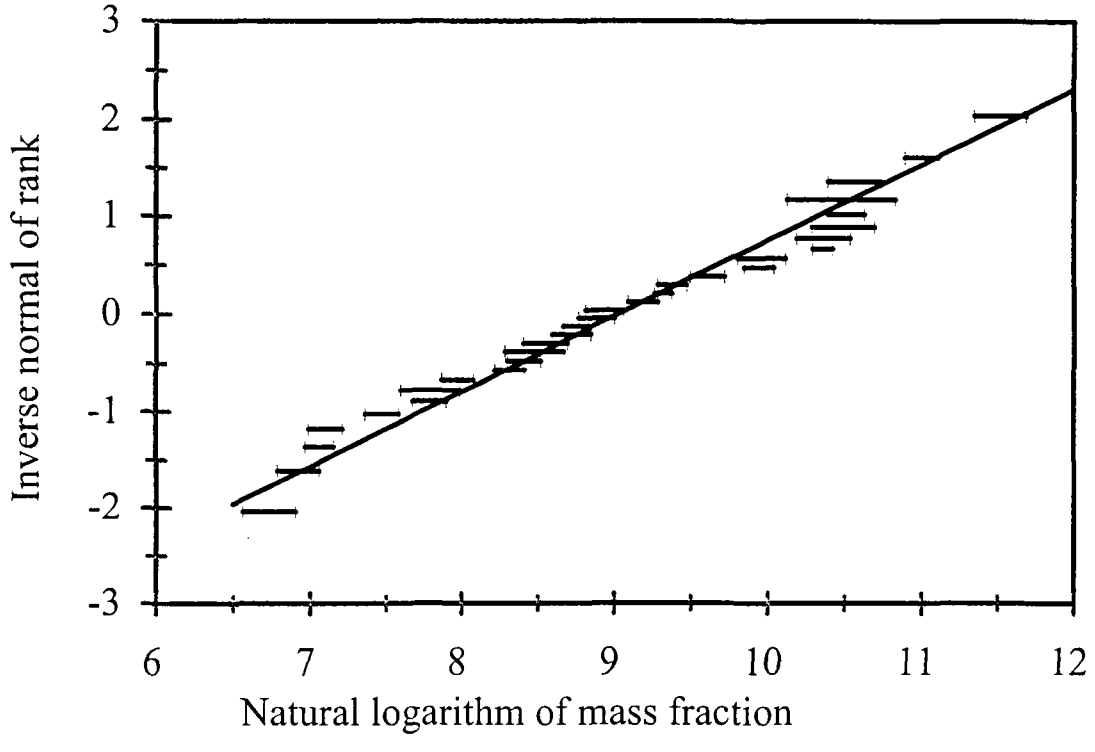

Figure 5.1 Probability plot for surface soil data from Plainwell former impoundment

For Otsego, the sample arithmetic mean and standard deviation of the sample average values are 13.7 mg/kg and 12.5 mg/kg respectively. Fitting a two-lognormal statistical distribution gives an MLE for the mean of 14.0 mg/kg, with the two components having means of 0.97 mg/kg and 17.9 mg/kg. The likelihood-based estimate of UCL95 on the mean is 21.9 mg/kg.

For Trowbridge, the sample arithmetic mean and standard deviation of the sample average values are 16.7 mg/kg and 19.0 mg/kg respectively. The statistical distribution model is the sum of three lognormals (three distinct parts of the distribution are clearly apparent in Figure 5.3), although the lognormal component with lowest mean degenerates to a constant at the maximum likelihood estimate, and was left constant for the rest of the analysis (this has negligible effect on the results). The statistical model gives an MLE for the mean of 19.6 mg/kg, and a likelihood-based estimate for the UCL95 on the mean of 29.3 mg/kg.

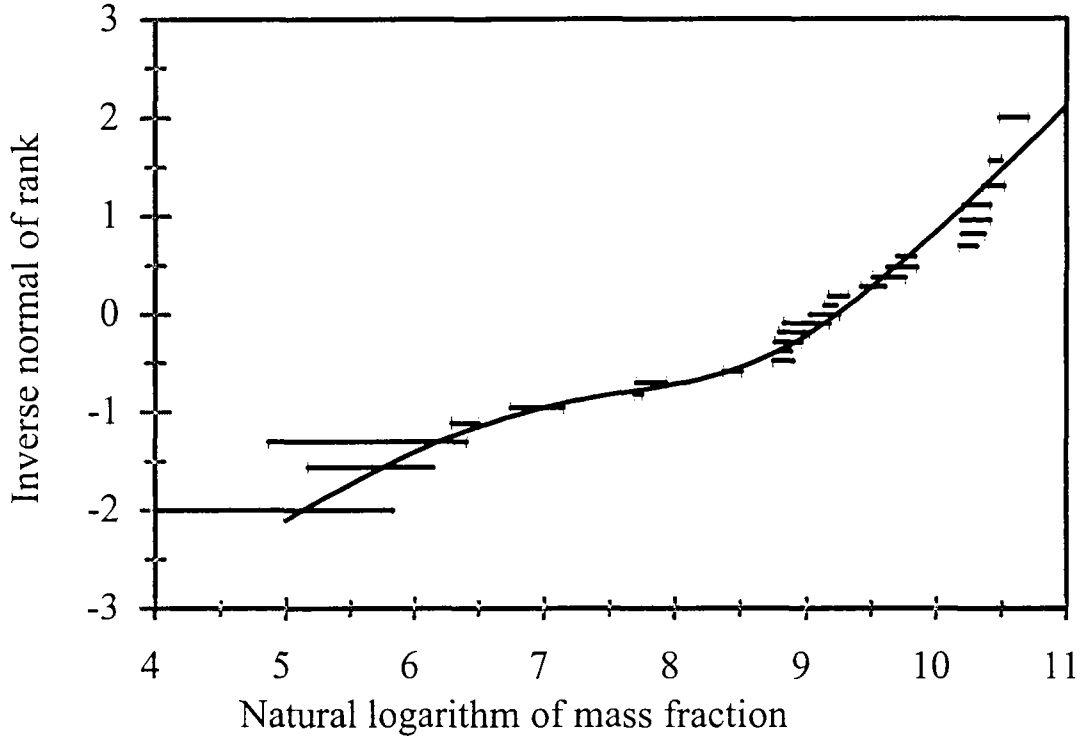

Figure 5.2 Probability plot for surface soil data from Otsego former impoundment

There is some evidence in the probability plots for three-lognormal statistical models in all cases, although the evidence becomes weaker from Trowbridge, to Otsego, to Plainwell — *i.e.* going upstream. The upstream gradient is suggestive of some physical basis for the three-distribution model; however, the fits obtained above with fewer components for Plainwell and Otsego are adequate. For comparison, a three-lognormal fit to the Otsego data give an MLE estimate of mean of 13.6 mg/kg, with a UCL95 estimate of 17.7 mg/kg, not substantially different from the two-lognormal fit.

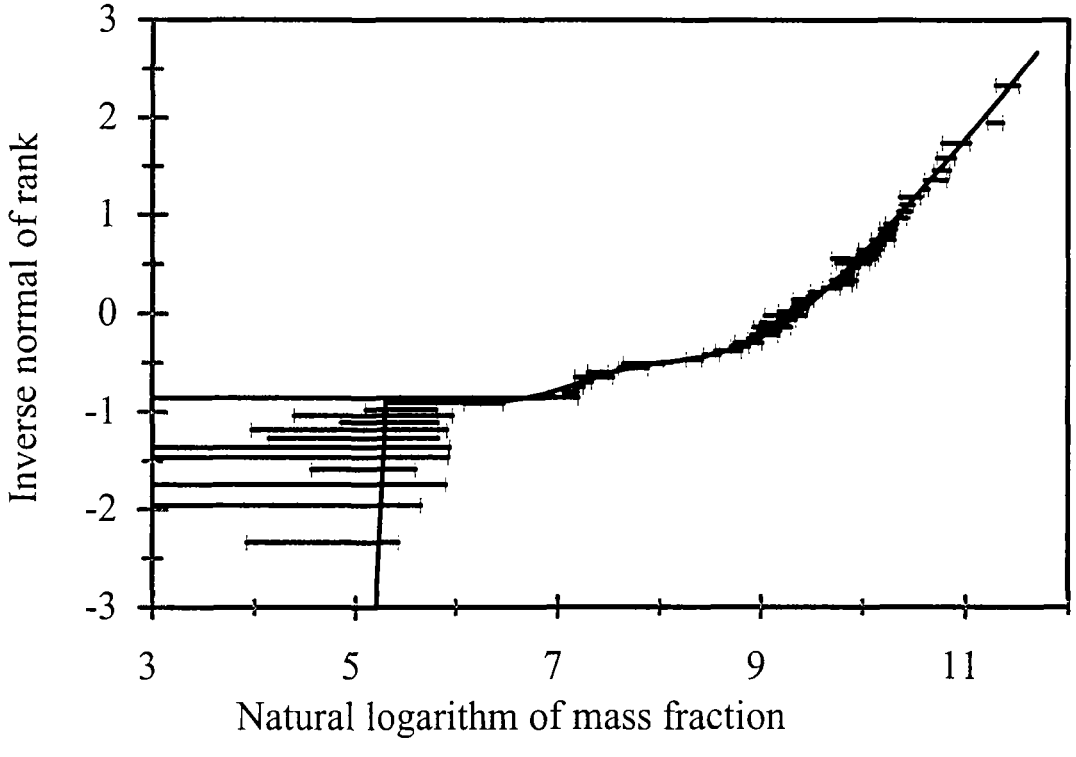

Figure 5.3 Probability plot for surface soil data from Trowbridge former impoundment

5.2 *The hunter/fisher scenario*

The hunter/fisher scenario considers the contact of a hunter or fisher with the PCB contaminated soils of the former impoundments, and the resulting exposures to PCBs. The approach taken is to obtain highly conservative (bounding) estimates of doses by deliberately choosing overestimates for several parameters. Dermal contact is examined in Section 5.2.1, and soil ingestion in Section 5.2.2. All the calculations are performed in the spreadsheet Other_exposure.wb3 (Appendix B.11). The scenario corresponds to a hunter or fisher staying on the former impoundment for a day's activity, 20 times per year (during the hunting season for a hunter; see below for the basis for this period).

5.2.1 *Dermal contact*

During each day's activity, the hunter/fisher incidentally comes into contact with the surface soil containing PCBs in the former impoundment area ("regular" events). In addition, every other day the hunter or fisher gets his/her hands muddy ("muddy hand" events) through activities such as moving aside the undergrowth, or securing waterfowl he or she has shot. Finally, once a year, the hunter/fisher ventures too far into a muddy area, and gets his or her feet muddy and perhaps loses a shoe ("muddy feet" events). Detailed justifications for the parameter values used for each of these events are presented below.

The dose rate from dermal contact with soil is obtained from:

$$D_{ad} = \sum_{event\ types} \frac{D_{a,event} E_v E_f E_d}{W_b T_a} \tag{5.4}$$

where the terms and their dimensions are:

D_{ad}	Dose rate, mass per unit body weight per unit time, averaged over the averaging period (T^{-1})
$D_{a,event}$	Absorbed mass per event (M)
E_v	Event rate during exposure (T^{-1})
E_f	Fraction of time exposed during the exposure duration (dimensionless)
E_d	Exposure duration (T)
W_b	Body weight (M)
T_a	Averaging period (T)

The absorbed mass per event, $D_{a,event}$, is obtained from

$$D_{a,event} = m \sum_i S_{a,i} w_i f_i \tag{5.5}$$

where the summation is over the affected body parts (labeled by i), and the terms and their dimensions are:

$S_{a,i}$	Surface area of body part i (L^2)
m	Mass fraction of PCBs in soil (dimensionless)
w_i	Skin loading of soil for this event (ML^{-2})
f_i	Fraction of PCBs absorbed (dimensionless).

The values used for these factors are:

E_v Event rate during exposure (T^{-1})

The event rate was set at 1/day, to match this scenario to the measurements used for estimating skin loading of soil.

E_f Fraction of time exposed during the exposure duration (dimensionless)

The fraction of time exposed was taken to be 20 days/year for "regular" events. This is a conservative upper end estimate of the time that a hunter might locate on the impoundments, given the available data on the average number of days spent hunting for various small game animals, and the relatively small size of the impoundments. The average over the periods 1992–1996 for the number of days spent hunting for various small game are (Karasek, 1998):

Game animal	Avg days in field
Pheasants	4.2
Quail	4.9
Ruffed Grouse	7.6
Woodcock	7.0
Ducks	7.0
Geese (Regular season)	6.5
Geese (Early season)	3.7
Geese (Late season)	4.0
Cottontail Rabbits	9.4
Snowshoe Hares	8.4
Squirrels	7.9
Crow	6.4

and the highest annual average during any one year for any of these game was 11.1 days/year (cottontail rabbits in 1993).

For fishers, individuals may fish more often than 20 days/year on average. Approximately 50% of those surveyed fished more than 6 times in the year before the Kalamazoo River Angler

Survey (MiCPHA, 2000a), and the average number of days fishing was 22 per year for active anglers in the Kalamazoo River basin (Atkin, 1994). However, it is highly unlikely that an angler would fish entirely from the banks of the former impoundments, because of the difficulty of reaching these areas, the uninviting nature of the banks at most shoreline locations in the impoundments, and the availability of many more accessible locations. Thus 20 days/year is a conservative overestimate for the average days/year spent by fishers on the former impoundments.

The "muddy hands" event was assumed to occur every other day. This sort of event corresponds to handling reeds, as might occur during construction of a blind, or perhaps to handling waterfowl or fish that have been dragged through mud.

The "muddy feet" events were taken to occur once per year, with the assumption that such events would be sufficiently unpleasant that they would be avoided as much as possible.

E_d Exposure duration (T)

The effective exposure duration was taken to be 40 years. This includes any period beyond actual exposure to account for the gradual decay of PCB concentrations in human tissues. There is no available information on the length of time that people might go wildfowl hunting on the Kalamazoo. The MDCH angler survey (MiCPHA, 2000a) has no direct information on the duration of fishing; however, for those who ate the fish, the average years of eating Kalamazoo River/Portage Creek fish was 10 to 12 years. For fishers in the Kalamazoo basin area, the average time spent fishing in Michigan was about 28 years (Atkin, 1994).

W_b Body weight (M)

Body weight was chosen to be the standard estimate of 70 kg.

T_a Averaging period (T)

The averaging period was taken to equal the exposure duration, for computation of non-cancer risks, and a standard lifetime of 70 years (corresponding to the definition used for cancer potency estimates) for cancer risk estimates.

$S_{a,i}$ Surface area of body part i (L^2)

The following areas were used for the various body parts:

Body part	Hands	Arms	Legs	Faces	Feet
Area (m^2)	0.099	0.291	0.64	0.13	0.131

These are the median surface areas for men (U.S. EPA, 1997, Table 6-2). The median estimate is chosen to correspond with the standard body weight of 70 kg — surface area per unit body weight is likely to decrease with increasing body weight, and there is unlikely to be substantial variation between men and women of similar weight.

The methodology adopted is one of those recommended in U.S. EPA (1997) — measured values for soil accumulation on all the appendages are summed. The torso is considered unexposed in any of the activities for the hunter/fisher scenario — hunting would take place principally in fall, requiring appropriate dress for both warmth and protection while pushing through brush, and the long-term anglers examined here are unlikely to have consistently exposed torsos.

w_i Skin loading of soil for this event (ML^{-2})

Skin loadings per event were estimated from the measurements of Kissel *et al.* (1996) and Holmes *et al.* (1999), as also reported in U.S. EPA (1997). For "regular" events, the measurements of groundskeepers were considered appropriately conservative. Mean values for the skin loadings were obtained from the distributions implied by the reports cited. The measurements were all of people wearing their normal clothing for the activities concerned, and were referred to the bare surface area of the body part concerned. Thus no correction for assumed different fractions of the skin surface being exposed are appropriate.

It was assumed that each individual measurement of groundskeepers reported by Kissel *et al.* (1996) and Holmes *et al.* (1999) represented individual events, with the distribution of values equivalent to the differences that would occur for any individual during different events. Since each of the five sets of measurements was reported to have a distribution of values consistent with lognormal, all the groundskeeper measurements were accumulated to obtain a grand lognormal distribution for all groundskeepers for each body part. The mean values for that lognormal distribution were then used to estimate the long-term average soil loading (averaged over many events).

The accumulated distributions were obtained by convolving the reported distributions for each of the five sets of groundskeepers for each body part separately. Where no measurement was reported for a particular body part for a particular set of groundskeepers, the convolution was performed over just the sets that did provide that body part measurement. Where no standard deviation was reported for a particular body part for a particular set of groundskeepers, its square was estimated as the average of the variances over the other sets for that body part, weighted by their degrees of freedom. No standard deviations were reported for measurements on feet — their squares were estimated as the average over the other body parts of the within-set degree-of-freedom-weighted mean variances. The convolution was performed analytically using the logarithms of the measurements, since they are normally distributed. That is, for each set j of measurements we have a mean w_{ij} of the logarithm of skin loading for body part i (the logarithm of the reported median skin loading) and a within-set unbiased standard deviation estimate s_{ij} (the logarithm of the reported geometric standard deviation, estimated as just described if necessary),

together with the number of samples n_j within the set. Convolving these gives the following estimates for mean w_i and standard deviation s_i of the combined set:

$$w_i = \frac{\sum_j n_j w_{ij}}{\sum_j n_j}$$

$$s_i^2 = \frac{\sum_j \left(\left(n_j - 1\right) s_{ij}^2 + n_j \left(w_{ij} - w_i\right)^2 \right)}{\left(\sum_j n_j \right) - 1} \tag{5.6}$$

The estimated mean skin loading was then obtained by transforming back from the resulting estimates for the mean w_i and standard deviation s_i of the lognormal distribution, as

$$\exp\left(w_i + s_i^2/2\right) \tag{5.7}$$

Table 5.2 shows the original data, the estimated overall distribution, and the estimated mean values.

For the "muddy hand" events, assumed to take place every other day during exposure, an additional soil loading to the hands alone was assumed, corresponding to the values reported by Kissel *et al.* (1996) for measurements on the hands of reed gatherers. Once again, the reported values were assumed to correspond to individual events from a distribution common to all participants, and the mean value of the assumed lognormal distribution used. The geometric mean and geometric standard deviation for measurements on the hands of the four reed gatherers were 0.66 mg/cm^2 and 1.8, leading to the mean estimate of 0.78 mg/cm^2 used here for this type of event.

For the "muddy feet" events, the a skin loading to the feet corresponding to that of the reed gathers reported by Kissel *et al.* (1996) was added to the above exposures. One of the four reed gatherers lost a shoe during the activity measured by Kissel *et al.* (1996), so the possibility of shoe loss is incorporated in this distribution. Once again, the same approach as used for the other two cases was used. The geometric mean and geometric standard deviation for measurements on the feet of the four reed gatherers were 0.63 mg/cm^2 and 7.1, leading to the mean estimate of 4.30 mg/cm^2 used here for this type of event.

Table 5.2 Geometric mean (GM) and geometric standard deviation (GSD) of skin soil loading, in mg/cm^2, for various body parts (data from U.S. EPA, 1997), for groundskeepers						
Set	Number in set	Hands	Arms	Legs	Faces	Feet
		GM GSD	GM GSD	GM GSD	GM GSD	GM GSD
1	2	0.15 (-)	0.005 (-)		0.0021 (-)	0.018 (-)
2	5	0.098 2.1	0.0021 2.6	0.001 1.5	0.01 2	
3	7	0.03 2.3	0.0022 1.9	0.0009 1.8	0.0044 2.6	0.004 (-)
4	7	0.045 1.9	0.014 1.8	0.0008 1.9	0.0026 1.6	0.018 (-)
5	8	0.032 1.7	0.022 2.8	0.001 1.4	0.0039 2.1	
Overall		0.046 2.29	0.0068 3.65	0.00092 1.63	0.0041 2.30	0.0093 2.74
Mean		0.0651	0.0158	0.00104	0.00581	0.0155

f_i Fraction of PCBs absorbed (dimensionless).

The fraction of PCBs absorbed is taken to be 6% as an upper bound estimate for the high organic carbon content soil in the impoundment areas. This is the upper end of the range recommended by U.S. EPA (1992a). It was set equal for all body parts and for all soil mass loadings. The HHRA uses a value of 0.14 for the absorption of PCBs, citing that value as coming from a 1998 draft of a U.S. EPA guidance document that is in internal review. Examination of the review draft shows that the value 0.14 is explicitly selected as an upper bound screening value. That value comes from an experiment (Wester *et al.*, 1993) in which soil of low organic carbon content (0.9%) was sieved to remove all fine materials (leaving sand that did not pass through #80 mesh) and then experimentally augmented with 67 ppm total Aroclors. The resulting sand, at a loading of approximately 40 mg/cm^2, was held against the skin of monkeys using modified human eye patches in such a way that the sand could move and rub against the skin for 24 hours. The results from such an experiment have no relevance to the potential absorption from high

organic carbon content soil (5% to 20% organic carbon) from the impoundments, adhering to the skin (not pressed against it) in a film in which the individual particles are not free to rub against the skin surface.

m Mass fraction of PCBs in soil (dimensionless)

The exposure point mass fraction ("concentration") of PCBs in soil.

5.2.2 Soil ingestion

The dose rate from soil ingestion is given by:

$$D_{ad} = \frac{mI_s E_f E_d f_a}{W_b T_a} \tag{5.8}$$

where the terms and their dimensions are:

D_{ad}	Dose rate, mass per unit body weight per unit time, averaged over the averaging period (T^{-1})
I_s	Soil ingestion rate (MT^{-1})
f_a	Fraction absorbed (dimensionless)
m	Mass fraction of PCBs in soil (dimensionless)
E_f	Fraction of time exposed during the exposure duration (dimensionless)
E_d	Exposure duration (T)
W_b	Body weight (M)
T_a	Averaging period (T)

The last five parameters take the same values as discussed under "Dermal contact," above.

I_s Soil ingestion rate (MT^{-1})

The (adult) soil ingestion rate is estimated to be about 50 mg/day (U.S. EPA, 1997). There are only two studies that have attempted to measure adult average soil ingestion rates (Calabrese *et al.*, 1990; and Stanek *et al.*, 1997), the former in six adults, the latter in ten. The U.S. EPA recommendation is based principally on Calabrese *et al.* (1990). The later study by Stanek *et al.* (1997) is said to "suggest lower levels of soil ingestion in adults than previous studies," based on 280 subject-days (10 subjects × 28 days) of evaluation, the largest amount of data available on soil ingestion in adults. Stanek *et al.* (1997) estimated that the average adult ingested 10 mg/day (although the uncertainty is large).

f_a Fraction absorbed (dimensionless)

The fraction of PCBs in ingested soil that are absorbed. Strictly speaking, what is required is a relative absorption fraction; the fraction absorbed from this soil, compared with the fraction absorbed in the studies used for determining the RfD and/or cancer potency factor. The value used here is 0.76, based on Fries *et al.* (1989) — the derivation of this value is discussed further in Appendix A.1.

5.2.3 Results for the hunter/fisher scenario

Applying the parameter values discussed above, at a soil concentration of 36.0 mg/kg, the upper 95th percentile confidence (UCL95) estimate for the average soil concentration at the Plainwell former impoundment, the average dose rate to the hunter/fisher during the period of exposure is 0.0024 µg/kg-day, which is substantially lower than the health-protective value of 0.05 µg/kg-day. The lifetime risk estimate is 2.8×10^{-6}, which is well within the range of U.S. EPA acceptable values, particularly since the populations involved are small. For the other two impoundments, UCL95 estimates of the average soil concentrations of PCBs for Otsego and Trowbridge are 21.9, and 29.3 mg/kg respectively. These lead to dose rates during exposure and risk estimates that are proportionately lower — dose rates during exposure of 0.0015 and 0.0020 µg/kg-day respectively, and lifetime risk estimate below 1.7×10^{-6} and 2.3×10^{-6} respectively.

5.3 The trespassing gardener scenario

This scenario has been included to evaluate the very few potential cases where someone trespasses on the former impoundments and sets up a vegetable garden in them. Only one such garden has been observed, and they contravene State Land Rules, so the expected population exposed is likely to be very small — almost certainly less than 5 people at any time. The same analysis as applied above for the hunter/fisher scenario was applied to the trespassing gardener scenario, with the addition of an estimate of doses from eating the vegetables grown in the garden. The observed garden has apparently has operated for up to 20 years (based on a contact with the gardener — personal communication, 2/2/2001, Laura Green, Cambridge Environmental Inc. to Brian von Gunten, Michigan Dept. of Environmental Quality). We here evaluate dermal exposure the soil of the impoundments, soil ingestion resulting from working on the soil, and ingestion of vegetables from the garden. All calculations are performed in the spreadsheet Other_exposures.wb3 (Appendix B.11).

5.3.1 PCB concentrations in garden soil and vegetables

PCB concentrations were obtained for five surface samples of soil from the one observed vegetable garden (CDM, 2000). Since cultivation of the soil is likely to change the PCB concentration of surface soils, by admixture with deeper soils and with any amendments applied to the garden, these measurements were considered more representative of the concentrations to which a gardener would be exposed than the measurements of unmodified surface soils from the rest of the impoundments. As for other soil samples, the results were reported as a best estimate of a mixture of Aroclors (Table 5.3), although many of the samples showed deviations from the congener patterns expected for such Aroclor mixtures. The reported best estimate Aroclor mixtures contained no Aroclor 1221, 1232, 1016, or 1248, so those Aroclors were treated as absent (1016 and 1248, while they occurred in impoundment surface samples, were rare there). For the other Aroclors, non-detects were treated as present at half the detection limit to obtain the total PCB concentrations shown.

Table 5.3 PCB concentrations in soil in the garden in the Otsego Impoundment

Sample Date	Sample name	1016	1221	1232	1242	1248	1254	1260	Total
		Concentration, mg/kg dry weight (< indicates practical quantitation limit)							
07/25/00	Gard 1-2	<0.0863	<0.0863	<0.0863	0.107	<0.0863	0.338	0.215	0.66
07/19/00	RRA-5	<0.0637	<0.0637	<0.0637	0.72	<0.0637	2.21	0.398	3.33
07/25/00	Gard 1-1	<0.145	<0.145	<0.145	0.899	<0.145	2.58	0.557	4.04
07/19/00	RRA-7	<0.545	<0.545	<0.545	1.36	<0.545	2.54	<0.545	4.17
07/19/00	RRA-8	<1.22	<1.22	<1.22	12.8	<1.22	3.88	<1.22	17.3

These concentrations are consistent with coming from a lognormal distribution, but the UCL95 estimate for the mean is outside the range of measurements (and also much higher than the estimates for mean concentration in the impoundment surface soil), so the estimate of soil concentration used is the maximum measured value of 17.3 mg/kg.

PCB concentrations were measured in eight samples of vegetables from the garden (CDM, 2000), and again reported as Aroclor mixtures; although again the congener patterns often did not match well. Aroclors 1016, 1221, 1232, and 1260 were reported as absent from all samples, with detection limits generally 0.0025 mg/kg wet weight (although approximately double that in potatoes). Table 5.4 lists detected Aroclors, with total PCB concentrations in the samples estimated by assigning half the detection limit to any undetected Aroclor that was reported in any produce sample.

Table 5.4	PCB concentrations in produce, and mean produce consumption by home gardeners in the Midwest.[a]					
Produce	1242	1248	1254	Total	Mean consumption	
	PCB concentration, mg/kg wet weight (< indicates practical quantitation limit [a]				g/kg-day	Table in U.S. EPA (1997)
Green Tomatoes	<0.00250	<0.00250	<0.00250	0.0038	1.18	13-59
Potato	0.00318	<0.00250	<0.00250	0.0057	1.77	13-60
Rhubarb	0.00446	<0.00250	0.00484	0.011		
Horseradish	<0.00251	0.00662	0.00269	0.011		
Peppers	0.00398	<0.00252	0.021	0.026	0.234	13-53
Cucumber	<0.00267	0.0236	<0.00267	0.026	1.00	13-42
Carrots	<0.00250	0.0272	0.0143	0.043	0.457	13-40
Lettuce	0.0146	<0.00454	0.0546	0.071	0.383	13-45
Average				0.019		

[a] Aroclors 1016, 1221, 1232, and 1260 were not detected in any produce samples.

The average concentration in produce eaten by an individual was estimated by weighting the estimated total PCB concentration in each measured vegetable with the mean "consumer-only" consumption of home-produced food items in the Midwest U.S., as measured in the National Food Consumption Survey (U.S. EPA, 1997, Chapter 13). The absence of any estimates for consumption for rhubarb and horseradish should not lead to an underestimate of average concentration, since these had concentrations below the estimated overall average.

This estimate of concentration uses the only available measurements in produce from a garden within the former impoundment areas. It is thus not obtained as a statistical upper-bound estimate on a series of measurements for such concentrations, as would be preferred. The produce samples were washed with regular water and blended in a food mixer before submission for analysis (personal communication, 2/2/2001, Laura Green, Cambridge Environmental Inc. to Brian von Gunten, Michigan Dept. of Environmental Quality). While such washing would remove surface soil, it does not entirely correspond to the effect of food preparation that might typically include removal of skin for potatoes and carrots, and discarding of less tasty parts of some of the other vegetables, and, for some vegetables, cooking. Skin removal might reduce the concentrations in the as-eaten vegetables, since PCBs are likely to accumulate more in the skin

and sub-surface parts of vegetables. Moreover, cooking of the vegetables might also reduce PCB concentrations.

The procedure of weighting the various concentration measurements by the mean consumption rates of consumers of individual home-grown vegetables is not strictly justified for such a small population of exposed people, since different individuals consume different vegetables in different proportions. However, no direct measurements are available on the actual population eating these vegetables, so these data are the best surrogates available.

5.3.2 PCB exposures for the trespassing gardener

The gardener was assumed to tend the garden for 100 days/year, and to have an effective exposure period of 25 years (taking account also of the persistence of PCBs in the body — see Section 6.3.2). The exposure from soil ingestion was calculated as for the hunter/fisher, with a soil concentration of 17.3 mg/kg, a fraction of time exposed of 100 days/year, an exposure duration of 25 years, and other parameters as indicated in Section 5.2.2. Dermal exposure was also calculated as for the hunter/fisher (Section 5.2.1), using a fraction of time exposed of 100 days/year, an exposure duration of 25 years, a soil concentration of 17.3 mg/kg, and the skin loading of soil as described in the next paragraph. Other parameters were as discussed in Section 5.2.1).

The skin loading of soil for a gardener was estimated from the measurements of Kissel *et al.* (1996) and Holmes *et al.* (1999), as also reported in U.S. EPA (1997). Of the persons measured, the farmers appeared the best surrogates for a gardener. Measurements on the two sets of farmers were combined to obtain estimates for soil loading — Table 5.5, calculated as for Table 5.2 in Section 5.2.1. Since no measurements were available for loadings on the feet, the value from groundskeepers (Table 5.2) were judged the nearest surrogate of those available.

Table 5.5	Geometric mean (GM) and geometric standard deviation (GSD) of skin soil loading, in mg/cm^2, for various body parts (data from U.S. EPA, 1997), for farmers; and resultant estimated mean concentration.					
Set	Number in set	Hands	Arms	Legs	Faces	Feet
		GM GSD	GM GSD	GM GSD	GM GSD	GM GSD
1	4	0.41 1.6	0.059 3.2	0.0058 2.7	0.018 1.4	
2	6	0.47 1.4	0.13 2.2	0.037 2.9	0.041 3	
Overall		0.445 1.46	0.095 2.67	0.018 4.52	0.029 2.57	0.0093 2.74
Mean		0.478	0.153	0.055	0.046	0.0155

PCB exposure from produce consumption was calculated as

$$D_{ad} = \frac{m_w I_p E_d}{T_a} \tag{5.9}$$

where the terms and their dimensions are:

D_{ad}	Dose rate, mass per unit body weight per unit time, averaged over the averaging period (T^{-1})
I_p	Total garden produce consumption, as a fraction of body weight (T^{-1})
m_p	Average mass fraction of PCBs in produce (dimensionless)
E_d	Exposure duration (T)
T_a	Averaging period (T)

An upper end estimate of total consumption of produce from the vegetable garden by an individual was estimated from the National Food Consumption Survey by using the 95th percentile seasonally adjusted total Midwest consumer-only total homegrown vegetable intake of 7.41 g/kg-day (U.S. EPA, 1997, Table 13-33). This corresponds to approximately 1.1 lbs/day of vegetables per person from the garden. The average mass fraction in produce is calculated in Section 5.3.1, the exposure duration is 25 years, and the averaging period 25 years for calculation of dose rate during exposure, and a standard lifetime of 70 years for calculation of lifetime average dose rate.

5.3.3 Results for the trespassing gardener

With the parameters discussed, the estimated average daily dose rate during exposure for the trespassing gardener is 0.15 μg/kg-day, approximately three times the 0.05 μg/kg-day Michigan health protective value. Taking account of the effective exposure period, the upper bound lifetime risk estimate is 10×10^{-5}, within the acceptable range of values for the U.S. EPA, but higher by a factor of about 10 than Michigan's limit for waste sites.

6 *Fish ingestion*

6.1 Method of evaluation for risk and hazard index

Fish ingestion is likely to be the route leading to greatest potential individual and population exposures to PCBs at the Kalamazoo River site. The population of fish-eaters on the Kalamazoo River was examined in Phase I of the Kalamazoo River Angler Survey (MiCPHA, 2000a,b). That survey identified and characterized the angler population utilizing the affected portions of the river, and determined fish consumption patterns. The object here is to estimate the distribution of lifetime average intake rates and risks for the population of anglers who eat the fish that they catch, and also to estimate the uncertainties in that distribution. That distribution is assumed to apply for the whole population (including the anglers and others to whom they provide fish) who eat fish from the relevant stretch of the Kalamazoo. For the purposes of risk assessment, most interest centers on the upper end of the distribution.

For any individual, the average intake I during exposure (fraction of body weight per unit time) of PCBs may be estimated from:

$$I = \sum_i C_i(\tau, t)\frac{mnf_i s}{M} \tag{6.1}$$

where the summation is over fish species, labeled by i ; while the effective lifetime average intake I_{eff} of PCBs may be estimated from

$$I_{eff} = \sum_i C_i(\tau, t)\frac{mnf_i s}{M}\cdot\frac{t + t_e}{T} \tag{6.2}$$

where the terms and their dimensions are:

I	Average intake rate of PCBs as a fraction of body weight per unit time during exposure (usually expressed in mg/kg-day) (T^{-1}),
I_{eff}	effective lifetime average intake rate of PCBs as a fraction of body weight per unit time (usually expressed in mg/kg-day) (T^{-1}),
τ	calendar time at which fish-eating starts for this individual (T), set to be 1999 in all the following analyses,
t	period for which fish-eating continues (T), see Section 6.3.1,

C_i average mass fraction (dimensionless)[6] in edible portions of fish species i over the time period τ to $\tau + t$, see Section 6.2,

m average mass of fish consumed per meal (M), see Section 6.5.3,

n average rate of eating fish meals (meals per unit time) (T^{-1}), see Section 6.5.1,

f_i average fraction (for this fisher) of the meals that are species i (dimensionless), see Section 6.5.2,

s fractional survival of PCBs through preparation and cooking methods used by this individual (dimensionless), see Section 6.6,

M average body weight over the period of eating fish (M), set to 70 kg,

T nominal lifetime of 70 years (T),

t_e effective additional exposure period due to persistence of PCBs in the body (T), see Section 6.3.2 and Section 6.4.

These equations give actual dose rate during exposure and an effective lifetime average dose rate, and represent the same models as used in the HHRA — the actual lifetime average dose rate is modified by the effective additional exposure period t_e to take account of the persistence of PCBs in the body (see Section 6.3.2). No additional term is incorporated for the fraction of PCBs that is absorbed in the gut, since it is assumed that PCBs in fish are well absorbed — strictly, the assumption is that PCBs in fish are absorbed as well as the PCBs were absorbed in the animal (rat) experiments that are the basis for the toxicity estimates.

For any individual, all the terms in these equations are uncertain (uncertainty). For different individuals, some of the terms will be different because of the differences between members of the population (variability). The evaluation performed here takes account of the variability and uncertainty by using Monte Carlo methods. That is, the uncertainty and variability distributions for each term are evaluated in the following sections, keeping track of any correlations between the various distributions. Then each equation is evaluated a large number of times with different samples from those distributions. The resulting sets of values for dose rates allow evaluation of the variability and uncertainty distributions for the dose rates.

The Monte Carlo algorithm for the combined uncertainty and variability analysis is the following:

Repeat a large number (5,000) of times {
    Choose a sample from the uncertainty distribution for each term on the right hand side of equations 6.1 and 6.2, taking account of correlations.
    Repeat a large number (50,000) of times {
        Choose a sample from the variability distribution of each term on the right hand side of equations 6.1 and 6.2, conditional (if necessary) on the

[6] As before, we distinguish between mass fraction and concentration where precise definitions are required, but in the text often use the term concentration to represent both.

values already obtained from the uncertainty distributions, and taking account of any correlations.

Calculate the corresponding sample value for each term on the right hand side of equations 6.1 and 6.2 using the uncertainty and variability sample values.

Calculate the average dose rates using the equations 6.1 and 6.2.

Store the calculated values.

} (end of the inner repetition)

From the stored values, construct the variability distributions for the average dose rates.

Calculate population averages from the variability distribution

Store the variability distribution (for example, store a set of percentiles of the distribution), and the averages.

} (end of outer repetition)

From the stored variability distributions for average dose rates, construct the uncertainty distribution for those distributions (for example, construct the uncertainty percentiles for each stored variability percentile), and for the stored population averages.

Calculate averages over the uncertainty distributions.

Print out the results in a convenient way and interpret them.

In this analysis, exactly this scheme was used. The inner (variability) loop was repeated 50,000 times, and 31 percentiles (at 0.5, 1, 2, 3, 4, 5, 7.5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 92.5, 95, 96, 97, 98, 99, 99.5 percent) were saved. The outer (uncertainty) loop was repeated 5,000 times (saving each of the 31 percentiles of the variability distribution each time, together with the arithmetic mean and standard deviation of the variability distribution, and the mean and standard deviation of the logarithms), and then 99 percentiles (at 1, 2, .. to 99 percent) of each of the 31 variability percentiles and the averages were calculated and saved.

For a randomly chosen individual, there is no difference between variability and uncertainty, because the randomness of choosing the individual transforms the variability between individuals into an uncertainty for the chosen individual. To evaluate the uncertainty distribution for a random individual, therefore, a simpler analysis may be performed:

Repeat a large number (1,000,000) of times {

Choose a sample from the uncertainty distribution for each term on the right hand side of equations 6.1 and 6.2, taking account of any correlations.

Choose a sample from the variability distribution of each term on the right hand side of equations 6.1 and 6.2, conditional (if necessary) on the values already obtained from the uncertainty distributions, and taking account of any correlations.

Calculate the corresponding sample value for each term on the right hand side of equations 6.1 and 6.2 using the uncertainty and variability sample values.

Calculate the average dose rates using the equations 6.1 and 6.2.
Store the calculated values.
} (end of repetition)
From the stored values, construct the uncertainty distributions for the average dose rates.
From the stored values, calculated average values and standard deviations.
Print out the results in a convenient way and interpret them.

This was the method used (using 1,000,000 repetitions), and the results were printed out at every 0.1 percentile point on the distribution. The complete program used to perform all calculations is described in Appendix C and provided in the supplemental electronic information.

The average dose rate during exposure obtained above may be compared with safe dose levels for non-cancer risks, and the lifetime average dose rate multiplied by a cancer potency estimate to estimate lifetime cancer risk. Alternatively, for the complete evaluation of the variabilities and uncertainties, we also take account of the variability and uncertainty of the carcinogenic potency and the minimum effective dose. The lifetime risk is thus estimated as

$$R = \sum_i \sum_j C_i(\tau, t)\beta_j \gamma_{ij} \frac{mnf_i s}{M} \cdot \frac{t + t_e}{T} \tag{6.3}$$

where

R is the lifetime risk estimate (dimensionless),
β_j is the carcinogenic potency for Aroclor j (1016, 1242, 1248, 1254, or 1260) (T), see Section 4.2,
γ_{ij} is the fraction of total PCBs that are Aroclor type j in fish species i (dimensionless), see Section 6.2.7,
and other terms have their previous meanings.

The variability and uncertainty distributions for the carcinogenic potencies for the Aroclors are discussed in Section 4.2, and the other terms are discussed in the following Sections.

For the ratio of average dose during exposure to a non-cancer minimum effect level for that particular exposure duration, we compute

$$H = \sum_i \sum_j C_i(\tau, t) \frac{\gamma_{ij}}{d(t)} \cdot \frac{mnf_i s}{M} \tag{6.4}$$

where

H is a hazard index (dimensionless),
$d(t)$ is the minimum effect level (T^{-1}) for Aroclors for an exposure duration t (equal to the period of fish-eating),
and other terms have their previous meanings.

The uncertainty and variability distributions for the minimum effect level, and how the minimum effect level varies with exposure duration, are discussed in Section 4.3; the value is not distinguished between Aroclors, but is applied to the total ingested dose of Aroclors.

The Monte Carlo algorithms for risk and hazard index are identical to those for dose, except that it is equations 6.3 and 6.4 that are evaluated, rather than equations 6.1 and 6.2.

Strictly speaking, several of the terms in these equations (for example, body weight, fish meal consumption rate) should depend on the age of the individual, and the whole equations should be averaged over the time of exposure rather than taking independent averages for each term. However, for the population examined here (fish-eating anglers who eat the fish they catch), the majority of the exposure for the more highly exposed individuals occurs during adulthood, where the age-dependent factors are fairly constant. Moreover, the major inter-individual variability in exposures is due to other factors — principally the variation in length of time during which fish eating occurs, and the rate of eating fish (number of fish meals per year). The population variability in these quantities is large enough to completely dominate the small inter-individual variabilities that might arise from age-dependent terms. Thus these equations are adequate to describe the distribution of an individual's effective lifetime average intake of, cancer risk from, and hazard index from PCBs.

The lifetime average intake rate I_{eff} is described as "effective" because it is supposed to represent the dose rate metric that is proportional to lifetime risk of cancer — it is slightly larger than the actual average intake, because of the incorporation of the effective extra exposure period t_e. The approach of averaging the dose rate over a lifetime, and the modification of incorporating a correction for the period that PCBs would remain in the body, in order to obtain a dose rate metric proportional to lifetime risk of cancer, must both be considered hypotheses that have not been directly tested. The former approach, however, is the current standard for performing risk assessments, so its uncertainty can be included in the overall uncertainty that has always to be linked with risk estimates based on standard cancer potency factors. The sensitivity to omission of the effective additional exposure period is examined in Section 6.11.1.

The following sections describe the individual terms in the equations for average intakes, cancer risk, and hazard index. Each relates how the term has been estimated from the available data for the site and for the population examined, and details the inter-individual variability and the uncertainty of those estimates as probability distributions (with correlations). Finally, the distribution of intakes of PCBs, the cancer risk estimates, and the hazard index estimates are obtained by using the Monte Carlo algorithms to combine all the separate variability and uncertainty distributions

.

6.2 *Concentrations in fish*

The HHRA (MiDEQ, 2000) states at page 3-19, Section 3.5.3, that

> Average and maximum concentrations were used to reflect a range of exposure point concentrations for the angler and nearby residents scenarios. These concentrations are presented on Tables 2-1 and 2-3. An attempt was made to calculate the upper 95 percent confidence limit (95% UCL) around the mean for both the fish and floodplain data sets. In both cases, the 95% UCL exceeded the maximum concentrations. As specified by USEPA guidance, the maximum concentration [*sic*] were therefore selected as the upper bound exposure point concentrations (USEPA, 1992)

The data from the former impoundments (part of what the HHRA considers the floodplain) have been evaluated in Section 5.1. The statements made about the fish concentrations are both false and misleading, and they omit an observation that is important for risk assessment.

First, they are misleading in that the maximum concentration in individual fish fillets is not an exposure point concentration for the purposes of long-term exposure estimates, as required in the HHRA and this assessment. No individual is exposed for his or her entire fishing lifetime to the concentration in just one fish. People will eat a series of fish, so they will be exposed to an average of the concentrations in many fish (or the concentrations in fillets from those fish, if the fish are filleted). If it is assumed that the concentration in fish is not changing with time, as is implicitly done in the HHRA, then the time average of the concentration in multiple fish eaten over a long time span (several years) would be adequately approximated by the average over multiple fish caught at a single time (or, in this case, caught at one or two distinct times). Thus the exposure point concentration required for the HHRA is the mean concentration in the sampled fish. To take some account of the uncertainty of that mean, it is usual (for the purposes of making a conservative point estimate of risk) to estimate an upper 95th percentile confidence limit on the mean.

Second, the statements are false. In no cases do the UCL95 (upper 95 percent confidence) estimates on the mean exceed the maximum concentrations using the fish data evaluated in the HHRA, as demonstrated in Section 6.2.1 below.

Third, the HHRA failed to take account of a very significant change in concentrations in the fish between the two times of the measurements that were used. Moreover, many more recent fish concentration data are available for the Kalamazoo River that should have been incorporated into the concentration estimates. The failure to incorporate more recent information from the site is especially noteworthy because the more recent fish measurements include more of the fish species that people actually eat.

6.2.1 UCL95 estimates for the HHRA dataset

The HHRA used the data from the 1993 and 1997 sampling episodes. It correctly used the fillet data to account for what people would eat, but the handling of non-detects was probably less than optimal (see the discussion of the soil samples in Section 5.1, where this issue was discussed). However, even with this handling of the non-detects the estimates of UCL95 for the means that are obtained are not higher than the maximum values.

The usual procedure used to estimate UCL95 is to first evaluate whether the distribution of values is consistent with lognormal. The best available method for this purpose is to use the Shapiro-Wilk statistic (Royston, 1982, 1993, 1995) in conjunction with probability plots. If the distribution is consistent with lognormal, then the procedure of Land (1971, 1973, 1974, 1975, 1988; Lyon & Land, 1999) is used to obtain an unbiased estimate of the UCL95 for the mean. If the distribution is not consistent with lognormal, the default approach is to use the t-statistic procedure that is an unbiased optimum procedure if the distribution is normal. It should be recognized, however, that the t-statistic procedure is robust against different distributions, and is asymptotically unbiased and optimum for almost all distributions (including lognormal). Land's procedure should not be used if the distribution is not consistent with lognormal, since the resulting estimate will be highly biased. Alternative statistical approaches may be warranted in particular circumstances — the discussion of the exposure point concentrations for the impoundments in this document illustrates such approaches.

Table 6.1 shows the analysis of the carp data used in the HHRA, with non-detects treated as in the HHRA.[7] In every case the distributions are consistent ($p>0.05$, Shapiro-Wilk statistic) with lognormal, and the estimates of UCL95 for the means are lower than the maximum measured values. The UCL95 estimates selected by the procedure described above are shown in **bold** font. Table 6.2 shows the same analysis for smallmouth bass. In every case except one, the distributions are consistent ($p>0.05$, Shapiro-Wilk statistic) with lognormal, and in every case the UCL95 estimates on the means are lower than the maximum measured values.

[7] The calculations for this section are in the spreadsheet Fish_data_HHRA.wb3, Appendix B.8.

Table 6.1	Total PCB concentrations in Carp, 1993 & 1997 combined, HHRA treatment of non-detects.					
ABSA [a]	Number of samples	Mean (ppm)	Shapiro-Wilk probability for lognormal	UCL95 estimates (ppm) [b]		
				Normal	Lognormal	Maximum measured value
1	22	0.09	0.24	0.10	**0.10**	0.27
2	22	0.43	0.91	0.59	**0.73**	1.9
3,4,5	44	5.7	0.31	6.7	**7.3**	17
6	11	3.5	0.29	4.8	**6.3**	8
7	11	2.7	0.69	3.7	**5.6**	6.4
8	11	4.6	0.67	6.1	**7.6**	9.6
9	22	1.2	0.96	1.7	**2.1**	6.5
10	11	7.6	0.50	10.4	**15.0**	17
11	23	4.8	0.68	6.3	**8.5**	17

[a] Aquatic Biota Sampling Area

[b] Estimates of the upper 95th percentile on the mean, assuming a normal or lognormal distribution, with maximum concentrations also indicated. The selected value (bold) is the lower of the maximum and the Lognormal value if the Shapiro-Wilk statistic indicates that the data are acceptably lognormal, otherwise the lower of the maximum and the Normal value.

Table 6.2 Total PCB concentrations in smallmouth bass, 1993 & 1997 combined, HHRA treatment of non-detects.						
ABSA	Number of samples	Mean (ppm)	Shapiro-Wilk probability for lognormal	UCL95 estimates (ppm) [a]		
				Normal	Lognormal	Maximum measured value
1	22	0.091	0.63	0.12	**0.13**	0.31
2	22	0.19	0.87	0.25	**0.32**	0.67
3,4,5	44	0.95	0.83	1.2	**1.3**	3.9
6	11	0.99	0.60	1.5	**1.8**	3.7
7	11	1.5	0.79	2.0	**2.5**	3.7
8	11	2.0	0.90	2.5	**2.7**	4.2
9	22	1.9	0.049	**2.5**	4.3	5.8
10	11	1.9	0.21	2.1	**2.2**	2.4
11	22	0.80	0.16	1.1	**1.1**	4.3

[a] Estimates of the upper 95th percentile on the mean, assuming a normal or lognormal distribution, with maximum concentrations also indicated. The selected value (bold) is the lower of the maximum and the Lognormal value if the Shapiro-Wilk statistic indicates that the data are acceptably lognormal, otherwise the lower of the maximum and the Normal value.

6.2.2 More complete statistical treatment of non-detects

The HHRA used data-summary statistics for total PCBs that were obtained by reporting the sum of detected Aroclors if any were detected, and the effective detection limit for an individual Aroclor if no Aroclors were detected[8] (the latter occurred only once each for carp and smallmouth bass data). Such a procedure may underestimate concentrations where only one or two Aroclors are detected at low concentrations — there are many samples where the estimated total Aroclor concentration is lower than the reported detection limit for the non-detected Aroclors — and may also distort the apparent distribution of concentration values.

[8] See also the discussion of the former impoundment soil data, in Section 5.1

The standard procedure is to initially examine all the data to determine the presence or absence of particular chemicals. Examination[9] of all the fish data (including whole fish data and the available 1999 sample data that were not used in the HHRA) shows that Aroclors 1221 and 1232 have never been detected in any sample, as also occurred for the transect soil samples. Although these Aroclors are mixtures of many individual congeners, the absence of their detection, together with the known more rapid environmental removal processes operating on the congeners making up these two Aroclors (which are the lowest chlorinated), indicate that these two Aroclors may be adequately treated as not present.

For the remaining Aroclors (1016, 1242, 1248, 1254, and 1260), it may be adequate to treat non-detects as ½ the detection limit without substantially biasing any estimates of concentrations. The analyses described below were carried out on the measured concentrations, using either the range of possible values (treating all non detects as zero or as equal to their detection limits to obtain this range), or using ½ the detection limit for non-detects when this procedure was found to be adequate.

6.2.3 Data evaluation

Extensive sampling of fish was carried out in 1993, 1997, and 1999, using essentially identical protocols; data from these three times are suitable for use in the risk assessment and are so used here. Other fish sampling has been performed, and this other sampling is used in the RI (BBL, 2000a) to evaluate time trends of PCB concentrations over a longer period. Since the time trends found in that evaluation are consistent with the trends present in just the 1993, 1997 and 1999 data, and the sensitivity of risk estimate results to the uncertainty of the time-trend is low (see Section 6.11.1), only the 1993, 1997, and 1999 data have been used here. This risk assessment is concerned with the Kalamazoo River from Morrow Dam to Lake Allegan Dam, so the measurements made in ABSAs [10] 11 and 13 are omitted from consideration. Measurements in ABSA 12 (Portage Creek) have been omitted also, because no fishing was observed in Portage Creek and no eating of fish was identified there. ABSA 10 is immediately downstream of the relevant stretch of river, so the measurements made there are used in the evaluation of the time trend but not for estimates of concentration. Summary estimates of measurements from ABSAs 1 and 2 have been included in some tables for comparison purposes, but these measurements are not used for the risk assessment. The sample measurements evaluated for the risk assessment are limited (as in the HHRA) to the parts of the fish that would be eaten — skin-off fillets for carp, catfish, and pike; skin-on fillets for smallmouth bass, walleye, sunfish, and bluegill; and whole body for suckers. The whole-body yearling bass measurements were not used in the risk assessment — these fish are too small at this age to be worthwhile food sources.

[9] See the spreadsheet Fish_data.wb3, Appendix B.10.

[10] Aquatic Biota Sampling Area.

Initial evaluation of the data using likelihood models that accounted for the range of values possible for each measurement (because of the non-detects for some Aroclors) showed that the results obtained were consistent with similar estimates made using ½ the detection limit (equivalent to the mid point of the range of potential values). The time-trend modeling described below retained the use of the range of values in a likelihood approach, but estimates of concentrations conditional on the time trend (as required for the risk assessment) were subsequently made using ½ the detection limit.

6.2.4 Time trend analysis

Time trends in PCB concentrations have been evaluated in the RI and supplement (BBL, 2000a,c) using sampling data on carp and smallmouth bass, some of which pre-date the 1993, 1997, and 1999 data used here. The RI analysis used a statistical model for the logarithm of total PCB concentration that accounted for fish length, weight, lipid content, and time, and found that PCB concentrations in fish have been decreasing at 5% to 10% per year.

A similar model for total PCB concentrations was applied to the 1993, 1997, and 1999 data for carp and smallmouth bass from ABSAs 3 through 10 (see spreadsheet Bass_Carp_time.wb3, Appendix B.9), after initial evaluations (Tables 6.3, 6.5, and 6.7) showed that at fixed sampling dates, and within each ABSA, the distributions of total PCB concentrations were consistent with being lognormal. The model applied may be written as:

$$\ln\left(C_{t,ijk}\right) = m_{ij} - \beta\left(t - t_0\right) + \lambda_j \ln\left(\frac{l_{ijk}}{\bar{l}}\right) + \delta_j \ln\left(\frac{f_{ijk}}{\bar{f}}\right) + \varepsilon_{ijk} \qquad (6.5)$$

where the subscripts indicate:

- i the ABSA (3 through 10),
- j the fish species (carp or smallmouth bass) within the ABSA,
- k the individual fish of the given species within the ABSA,

and the terms are:

- C_t total PCB mass fraction within the given fish (dimensionless),
- m mean logarithm of total PCB mass fraction for that species in that ABSA (dimensionless),
- t the calendar year of the measurement (T),
- t_0 a fixed calendar year, taken to be 1999 (T),
- β a decay constant indicating the average rate of removal of PCBs from availability to fish (T^{-1}),
- λ,δ coefficients indicating associations with weight, length, and lipid content, respectively (dimensionless),
- l,f fish length and lipid content of the sample, respectively, for individual fish; the overscore indicates the mean value over all measurements (L and dimensionless, respectively).

ε a random variate, normally distributed with mean zero and standard deviation dependent on ABSA and species (dimensionless).

This model was fit to the potential range (due to non-zero detection limits) of concentrations for each measurement using likelihood methods. Very similar results are obtained if ½ the detection limits are used to obtain point estimates. Preliminary analysis using the selected model was also extended to incorporate separate time trends for carp and bass. It showed that the time trends were not distinguishable (p=0.15, 1 degree of freedom, likelihood ratio test). Similarly, a preliminary analysis with a species-specific terms proportional to the logarithm of the individual fish weights showed that such terms were unnecessary (p=0.63, 2 degrees of freedom, likelihood ratio test) — the length and weight of these fish are highly correlated, with an approximate power law relation between them. Lastly, it is evident from summary statistics (Tables 6.4, 6.6, and 6.8) that the standard deviation within each ABSA does not materially depend on time.

The selected model was used to obtain the best estimates for the time trend, β, and its uncertainty distribution. The maximum likelihood estimate for the time trend was 0.0481 per year (about a 5% per year reduction in concentration), and the distribution obtained by the likelihood profile method was almost perfectly normal, with standard deviation 0.0134 per year. This time trend is similar to those found over the period extending back to 1983 (BBL, 2000a,c). It is not credible that the concentration of PCBs is increasing — indeed, there is a lower bound that can be placed on the long-term average rate of decrease. The total quantity of PCBs in the river sediments and former impoundments has been estimated at 53,800 kg (BBL, 2000), while the river water alone is removing approximately 30 kg/yr (see spreadsheet Other_exposures.wb3, Appendix B.11), based on measurements of PCBs in water (in year 2000) and river flow rate, giving a rate of decrease by water transport of approximately 5.5×10^{-4} per year. To conservatively account for possible underestimation of the quantity of PCBs in sediments or overestimation of water concentrations, we place a lower bound on the time trend of 1×10^{-4} per year.[11]

The time trend in average PCB concentrations in fish was estimated using data from carp and bass. For fish other than carp and bass, there are insufficient measurements over time to separately estimate time trends, or to include them in the analysis. However, the same time trend is expected to apply to all fish, since it measures the decrease in availability of PCBs which is common to all fish. Thus the same time trend was applied to all fish, and to turtles.

[11] Theoretically, in order for there to be any risk from the PCBs in the river, there must be some loss rate of PCBs from the river, otherwise no PCBs could get to people. It follows that the total effect of the PCBs over all time must be finite, since there is a finite quantity of PCBs in the river.

6.2.5 Uncertainty distributions for fish concentrations

The method described in Section 6.2.4 for time trend analysis could be applied to estimate the concentration of PCBs in carp and/or bass in each ABSA. However, that analysis took account of the length and lipid content of the fish, and so would correct for these. Fishers take whatever fish they find, not fish of standard length and lipid content, so using the model directly would underestimate the time-to-time variability of the concentrations in the fish. While that variability could be added back in, it is easier to use the data more directly in a way that can be applied to all the fish sampled.

Given the time trend, all the measurements on a given species in a particular ABSA can be reduced to an index date (here 1999, but the choice is immaterial) by modifying the measured concentrations by the change expected between the measurement date and the index date. The resulting concentration estimates will form a distribution that includes the effects of variation in lipid content, length, and any other factors except the time variation. The statistics of the distribution at the index date can be obtained from the statistics at each individual date and the time trend; thus the uncertainty distribution for mean and standard deviation of the logarithm of concentrations can be readily estimated, conditional on the time trend. Since the distribution of fish concentrations at any time is lognormal, this immediately allows estimation of the uncertainty distribution for the mean concentration of the fish at the index time, conditional on the time trend, and its average value over any particular time period.

Thus, given a decay rate β for the concentration (so the logarithm of concentration is decreasing linearly with time), the sample mean (M) and unbiased estimate of standard deviation (S) for the logarithms of the measurements corrected to the index time for fish of a given species within a given ABSA can be obtained. For if m_i and s_i are the sample mean and unbiased estimates of standard deviation for the logarithm of the n_i concentration measurements at time t_i before the index time, we get:

$$M = \frac{\sum_i n_i m_i - \beta \sum_i n_i t_i}{N}$$

$$(N-1)S^2 = \sum_i \left\{ (n_i - 1)s_i^2 + n_i \left(M - m_i + \beta t_i \right)^2 \right\} \qquad (6.6)$$

$$\text{where } N = \sum_i n_i$$

Then S^2/σ^2 is chi-squared distributed with N–1 degrees of freedom, where σ is the (unknown) true standard deviation of the logarithm of concentration in the index year; and $(M-\mu)\sqrt{N}/\sigma$ is independently normally distributed, where μ is the true mean of the logarithm of concentrations in the index year. An estimate, biased high, for the uncertainty distribution for the mean concentration in the index year, conditional on the time trend, can be obtained by sampling from these distributions to obtain independent samples s and m as explained in Appendix C.1.10, and

forming $\exp(m + \frac{1}{2} s^2)$. One additional constraint was applied — if the estimate so obtained for the mean concentration was more than 10 times higher than the highest concentration seen at any time, then that estimate was censored and another sample selected.[12]

The average concentration over times T_1 to T_2 after the index date, again conditional on the time trend β, is then:

$$\exp\left(m + s^2/2\right)\frac{\exp\left(-\beta T_1\right) - \exp\left(-\beta T_2\right)}{\beta\left(T_2 - T_1\right)} \tag{6.7}$$

This approach was followed using ½ detection limits for all non-detects, since the time-trend and similar analyses indicate that little or no bias is introduced by this approximation.

6.2.5.1 Carp

Some summary statistics (computed in spreadsheet Fish_data.wb3, Appendix B.10) for the sampling events (by year and ABSA sampled) are given in Table 6.3, using ½ the detection limit for non-detects. All but one (ABSA 1 in 1997) are consistent at a probability of 0.05 with lognormal distributions for the total PCBs in individual fish. UCL95 estimates are provided (**bold** font indicates the value that would be selected by the standard procedure already described), but these are not directly used here (they are just part of the distribution of values that are used).

Table 6.4 provides the summary statistics used for the risk assessment, as described in Sections 6.2.5 and 6.2.6 . These statistics are the sample mean and standard deviation of the logarithms of total PCB concentration, and the number of samples.

[12] This constraint was applied because of the very long tail of the sampling distribution for the mean of a lognormal when there are very few samples. In particular, without this constraint the turtle data, with only 6 measurements, gave many samples for the mean in the Monte Carlo procedure that were physically impossible. It is a more conservative version of the standard U.S. EPA procedure whereby if estimates of the average concentration exceed the maximum, the maximum concentration is used.

6.2.5.2 Smallmouth Bass

Summary statistics (computed in spreadsheet Fish_data.wb3, Appendix B.10) for each sampling event by year and ABSA are given in Table 6.5. Again, estimates of the UCL95 on the means are provided (**bold** font indicates the value that would be selected by the standard procedure already described), but these are not used here. Three of the sampling events provided data that are not consistent at p=0.05 with a lognormal distribution, but all are consistent at p=0.01. Of those used in this risk assessment, only ABSA 9 shows inconsistency at p=0.05. The inconsistency is sufficiently small (and within expectations considering the number of tests of lognormality performed) to maintain the assumption of lognormal distributions.

Table 6.6 provides the summary statistics used for the risk assessment, as described in Sections 6.2.5 and 6.2.6 . These statistics are the sample mean and standard deviation of the logarithms of total PCB concentration, and the number of samples.

Table 6.3 Summary statistics for Carp sampling, using ½ detection limit for non-detects								
Year	ABSA	Number of fish	Total PCBs (ppm)		Test [a] for log-normal	UCL95 estimates (ppm) [b]		
			Mean	SD		Normal	Lognorm	Max
93	1	11	0.16	0.04	0.056	0.18	**0.18**	0.24
93	2	11	0.72	0.61	0.94	1.1	**1.4**	2.3
93	3	11	5.0	2.4	0.19	6.2	**7.1**	9.0
93	4	11	7.4	3.8	0.22	9.5	**13.5**	14.1
93	5	11	6.3	4.7	0.69	8.9	**11.5**	18.7
93	6	11	3.8	2.7	0.37	5.3	**7.2**	8.5
93	7	11	3.0	1.9	0.83	4.1	**5.4**	6.8
93	8	11	5.1	3.2	0.69	6.8	**8.5**	11.1
93	9	11	1.9	1.8	0.32	2.9	**4.6**	6.9
93	10	11	8.2	5.7	0.70	11.3	**16.0**	19.2
97	1	11	0.18	0.06	0.009	**0.21**	0.21	0.34
97	2	11	0.39	0.2	0.58	0.5	**0.6**	0.9
97	5	11	6.3	6.0	0.25	9.6	**16.5**	18.3
97	9	11	0.86	0.6	0.28	1.2	**1.4**	2.0
99	2	11	0.62	0.4	0.13	0.8	**1.0**	1.3
99	4	11	6.9	5.6	0.22	9.9	**10.6**	22.7
99	5	11	10.4	3.1	0.70	12.2	**12.7**	17.4
99	8	11	3.0	1.9	0.065	4.0	6.4	**5.6**
99	9	11	1.9	1.8	0.11	2.9	6.1	**6.1**

[a] Shapiro-Wilk probability statistic evaluated for the logarithms of the total PCB concentrations. A value larger than 0.05 indicates a distribution that is consistent with lognormal.
[b] Upper 95th percentile estimates for the mean, assuming using the t-statistic (Normal) or Land's procedure (Lognorm), or the maximum among the measurements (Max). The value that would be selected by the standard procedure is shown in **bold**.

Table 6.4 Summary statistics for the logarithm of PCB concentrations in Carp (ABSAs 3 through 9).				
Year	ABSA	Number of samples	Sample mean of logarithms of concentration [a]	Sample standard deviation of logarithms of concentration [a]
93	3	11	1.491	0.504
93	4	11	1.833	0.687
99	4	11	1.743	0.598
93	5	11	1.621	0.715
97	5	11	1.411	0.978
99	5	11	2.305	0.306
93	6	11	1.112	0.739
93	7	11	0.899	0.700
93	8	11	1.439	0.649
99	8	11	0.831	0.821
93	9	11	0.298	0.910
97	9	11	-0.342	0.627
99	9	11	0.203	1.062

[a] Natural logarithm of the total PCB concentration measured in mg/kg, using ½ the detection limit for non-detects.

Table 6.5 Summary statistics for smallmouth bass sampling, using ½ detection limit for non-detects.								
Year	ABSA	Number of fish	Total PCBs (ppm)		Test [a] for log-normal	UCL95 estimates (ppm) [b]		
			Mean	SD		Normal	Lognorm	Max
93	1	11	0.21	0.077	0.034	**0.26**	0.26	0.38
93	2	11	0.37	0.20	0.18	0.47	**0.50**	0.82
93	3	11	1.2	0.89	0.65	1.7	**1.9**	3.6
93	4	11	0.54	0.17	0.079	0.63	**0.70**	0.77
93	5	11	2.0	0.87	0.52	2.4	**2.6**	4.1
93	6	11	1.1	1.0	0.39	1.6	**1.8**	3.9
93	7	11	1.6	1.1	0.83	2.2	**2.7**	4.5
93	8	11	2.2	1.0	0.87	2.7	**3.0**	4.4
93	9	11	3.7	1.6	0.61	4.6	**5.0**	6.6
97	1	11	0.15	0.018	0.43	0.16	**0.16**	0.18
97	2	11	0.20	0.083	0.16	0.24	**0.25**	0.42
97	5	11	0.56	0.42	0.79	0.79	**0.89**	1.7
97	9	11	0.61	0.44	0.23	0.85	**0.90**	1.8
99	2	11	0.34	0.35	0.012	**0.53**	0.52	1.4
99	4	11	0.78	0.26	0.29	0.92	**0.98**	1.1
99	5	11	0.58	0.34	0.064	0.76	**0.93**	1.2
99	6	11	1.2	0.84	0.26	1.6	**2.6**	2.7
99	8	11	0.83	0.33	0.77	1.0	**1.1**	1.5
99	9	10	0.65	0.24	0.038	**0.79**	0.93	0.99

[a] Shapiro-Wilk probability statistic evaluated for the logarithms of the total PCB concentrations. A value larger than 0.05 indicates a distribution that is consistent with lognormal.
[b] Upper 95th percentile estimates for the mean, assuming using the t-statistic (Normal) or Land's procedure (Lognorm), or the maximum among the measurements (Max). The value that would be selected by the standard procedure is shown in **bold**.

Table 6.6 Summary statistics for the logarithm of PCB concentrations in Bass (ABSAs 3 through 9).				
Year	ABSA	Number of samples	Sample mean of logarithms of concentration [a]	Sample standard deviation of logarithms of concentration [a]
93	3	11	0.020	0.596
93	4	11	-0.683	0.386
99	4	11	-0.307	0.353
93	5	11	0.588	0.430
97	5	11	-0.779	0.629
99	5	11	-0.706	0.607
93	6	11	-0.175	0.692
99	6	11	-0.134	0.841
93	7	11	0.296	0.653
93	8	11	0.685	0.448
99	8	11	-0.260	0.382
93	9	11	1.225	0.441
97	9	11	-0.660	0.555
99	9	10	-0.522	0.465

[a] Natural logarithm of the total PCB concentration measured in mg/kg, using ½ the detection limit for non-detects.

6.2.5.3 Other fish and turtles

Various other species of fish were sampled during 1993 and 1999. In 1993, suckers (white sucker, golden redhorse, northern hogsucker, white sucker, or spotted sucker) were sampled in ABSAs 1 through 12. In 1999, partly in response to the findings of the Kalamazoo River Angler Survey (MiCPHA, 2000a,b), various other fish were sampled — channel catfish in ABSAs 9 and 13, northern pike in ABSAs 9, 11, and 13, pumpkinseed sunfish in ABSA 4 and Bluegill sunfish in ABSAs 6, 8, and 9, walleye in ABSAs 9, 11, and 13, and yearling bass in ABSAs 2, 5, 8, 9, 11, and 13. Summary statistics for all sampling in ABSAs 1 through 9 are given in Table 6.7. The data for yearling bass are not used in the risk assessment — these fish are too small to be a significant part of any angler's fish consumption.

As for bass and carp, the distributions of concentrations of total PCBs in fish were generally consistent with lognormal statistics (see spreadsheet Fish_data.wb3, Appendix B.10). For ABSAs 3 through 9, the only potential exceptions were White Sucker in ABSAs 8 and 9 in 1993. However, the deviations are not large ($p>0.01$ in both cases), and are within the expected range given the number of comparisons performed here. Little bias should arise from treating the distributions of concentrations of total PCBs in fish as lognormal.

Snapping turtles were sampled in ABSA 5 in 1993, and in ABSAs 1 and 10 in 1994, and summary statistics for turtle muscle measurements in ABSAs 1 and 5 are provided in Table 6.7. The distribution of measurements in ABSA 1 is skewed by a single large measurement of 8.7 mg/kg — the other four measurements of turtle muscle were around 0.05 mg/kg total PCBs.

Table 6.8 provides the summary statistics (computed in spreadsheet Fish_data.wb3, Appendix B.10) used for the risk assessment, as described in Sections 6.2.5 and 6.2.6. These statistics are the sample mean and standard deviation of the logarithms of total PCB concentration, and the number of samples. For purposes of this risk assessment, all the suckers were considered equivalent, and all the sunfish were considered equivalent to panfish; this is the classification used in the Kalamazoo River Angler Survey (MiCPHA, 2000a,b) to evaluate fish consumption — see Section 6.5.2.

Table 6.7 Summary statistics for other fish, using ½ detection limit for non-detects.									
Species	Year	ABSA	Number of fish	Total PCBs (ppm)		Test [a] for log-normal	UCL95 estimates (ppm) [b]		
				Mean	SD		Normal	Lognorm	Max
White Sucker	93	1	11	0.16	0.03	0.17	0.17	**0.17**	0.22
Golden Redhorse	93	2	11	0.65	0.19	0.43	0.75	**0.80**	1.02
Northern Hogsucker	93	3	11	0.92	0.19	0.38	1.03	**1.06**	1.16
Golden Redhorse	93	4	11	2.50	0.50	0.18	2.77	**2.85**	3.18
Pumpkinseed Sunfish	99	4	11	0.43	0.15	0.62	0.51	**0.54**	0.75
Golden Redhorse	93	5	11	2.40	0.62	0.06	2.74	**2.90**	3.48
Golden Redhorse	93	6	11	2.31	1.02	0.16	2.87	**3.22**	4.85
Bluegill Sunfish	99	6	11	0.37	0.23	0.48	0.49	**0.54**	0.92
Golden Redhorse	93	7	11	2.30	0.45	0.73	2.55	**2.60**	3.07
White Sucker	93	8	11	0.91	0.34	0.040	**1.10**	1.29	1.28
Bluegill Sunfish	99	8	11	0.40	0.18	0.15	0.49	**0.58**	0.70
White Sucker	93	9	11	0.88	0.42	0.010	**1.11**	1.13	1.85
Bluegill Sunfish	99	9	8	0.49	0.18	0.98	0.61	**0.67**	0.85
Channel Catfish	99	9	11	1.34	1.00	0.96	1.89	**2.45**	3.60
Northern Pike	99	9	11	2.07	1.45	0.90	2.87	**4.01**	5.58
Walleye	99	9	11	0.86	0.48	0.75	1.12	**1.43**	1.68
Yearling Bass	99	2	5	0.69	0.18	0.18	0.86	**0.90**	1.00
Yearling Bass	99	5	5	1.93	0.47	0.96	2.38	**2.56**	2.62
Yearling Bass	99	8	5	2.30	0.44	0.63	2.72	2.89	**2.88**
Yearling Bass	99	9	4	1.29	0.53	0.18	1.91	4.24	**1.79**
Turtle	94	1	6	1.49	3.53	0.0001	**4.40**	2121	8.7
Turtle	93	5	6	0.78	0.69	0.094	1.34	2.90	**1.92**

[a] Shapiro-Wilk probability statistic evaluated for the logarithms of the total PCB concentrations. A value larger than 0.05 indicates a distribution that is consistent with lognormal.

[b] Upper 95th percentile estimates for the mean, assuming using the t-statistic (Normal) or Land's procedure (Lognorm), or the maximum among the measurements (Max). The value that would be selected by the standard procedure is shown in **bold**.

Table 6.8 Summary statistics for the logarithm of PCB concentrations in other fish and turtles (ABSAs 3 through 9).

Species	Year	ABSA	Number of samples	Sample mean of logarithms of concentration [a]	Sample standard deviation of logarithms of concentration [a]
Northern Hogsucker	93	3	11	-0.104	0.225
Golden Redhorse	93	4	11	0.897	0.217
Pumpkinseed Sunfish	99	4	11	-0.898	0.344
Golden Redhorse	93	5	11	0.838	0.298
Golden Redhorse	93	6	11	0.747	0.465
Bluegill Sunfish	99	6	11	-1.146	0.542
Golden Redhorse	93	7	11	0.815	0.203
White Sucker	93	8	11	-0.174	0.468
Bluegill Sunfish	99	8	11	-1.032	0.516
White Sucker	93	9	11	-0.209	0.395
Bluegill Sunfish	99	9	8	-0.772	0.367
Channel Catfish	99	9	11	0.057	0.726
Northern Pike	99	9	11	0.495	0.755
Walleye	99	9	11	-0.320	0.637
Turtle	93	5	6	-0.553	0.815

[a] Natural logarithm of the total PCB concentration measured in mg/kg, using ½ the detection limit for non-detects.

6.2.6 *Variability distribution for fish concentrations*

Of the 352 anglers in Phase I of the Kalamazoo River Angler Survey (MiCPHA, 2000a,b) who indicated that they eat fish or turtles (a yes response to Question 5, or a non-zero number of years eating fish), 60 were in Kalamazoo County and 292 in Allegan county (see spreadsheet Phase_1.wb3, Appendix B.12). Since individuals may fish from the same location for long periods, some account has to be taken of the variation in PCB concentrations in fish in different stretches of the river — this contributes to the variability between individuals. The variation was accounted for by selecting concentrations estimated in ABSA 3 or 4 (Kalamazoo County) with probability 0.17 (=60/352), and ABSAs 5 through 9 (Allegan County, above the Allegan Dam) with probability 0.83. Within each county, the ABSA was selected at random with equal probability. If concentration data for a particular fish type were not available for a selected ABSA, the nearest ABSA upstream or downstream (with equal probability) for which a concentration estimate is available was selected (ignoring county boundaries). If there was no available estimate in the selected direction, then the other direction was chosen. This procedure was performed independently for the different fish types considered.

6.2.7 *Aroclor fractions in fish*

The Aroclor fractions in the various fish vary slightly. Since different Aroclors are assigned different carcinogenic potencies in Sections 4.2.2, 4.2.3, and 4.2.4, and people eat different fish, this variability induces a variability in the population risk estimates. To take this into account, the Aroclor fractions in the various fish in ABSAs 3 through 9 were estimated by averaging over all the fish included in the analysis. For this averaging, non-detects were treated as half the detection limit, with Aroclors 1221 and 1232 treated as absent (zero concentration), and the average was obtained simply by treating all measurements with equal weight. The average Aroclor fractions are shown in Table 6.9 (see spreadsheet Fish_data.wb3, Appendix B.10).

Table 6.9 Aroclor fractions in the various fish types.					
Aroclor	1016	1242	1248	1254	1260
	Fraction of total PCBs				
Carp	0.057	0.078	0.317	0.474	0.074
Bass	0.112	0.123	0.133	0.547	0.085
Catfish	0.067	0.124	0.095	0.486	0.228
Panfish	0.074	0.241	0.165	0.442	0.078
Pike	0.044	0.114	0.067	0.686	0.088
Sucker	0.097	0.107	0.228	0.512	0.057
Walleye	0.052	0.128	0.297	0.455	0.069
Turtle	0.029	0.029	0.029	0.029	0.882

6.3 Exposure period

The HHRA indicated (Table 3-3) an exposure period of 30 years, with no reference provided, and intended to add 9 years to that exposure period to account for continued circulation of the PCBs in the bloodstream after cessation of exposure. The required actual exposure period for fish eaters is the length of time for which they eat fish during a lifetime. In addition, an effective additional exposure period may be required to account for the accumulated PCBs — however, this additional period must currently be considered hypothetical.

6.3.1 Actual exposure period for fish-eating anglers

Information on the length of time that the population of current anglers had eaten fish from the Kalamazoo up to the time of the survey was obtained in Phase I of the Kalamazoo River Angler Survey (MiCPHA, 2000a,b). These data provide a variability (between individual) distribution of time-to-present that current fish eaters had eaten fish, and may be used to estimate the lifetime length of time eating fish (Israeli and Nelson, 1992; the spreadsheet Phase_1.wb3, Appendix B.12 contains the analyses discussed in this section). Since the population studied is exactly the population of interest, these data provide the best available estimate for the population of Kalamazoo fish eaters. The potential bias from differential probability of inclusion in the Phase I survey due to different frequency of fishing has been ignored, because the probability differences are not too large (see Section 6.7), and the correlation appears to be relatively small (spreadsheet Meals.wb3, Appendix B.15, shows that the Pearson correlation coefficient between exposure period and the reported number of times fishing in the last calendar year is small, about 0.13).

Table 6.10 shows the distribution of responses to the question "How many years have you been eating fish you caught from the Kalamazoo River or Portage Creek" among the 286 respondents who provided estimates.[13] As can be seen, there are very few responses at longer times. For this reason, and because the question asked does not directly obtain the distribution required, the empirical data were fitted by theoretical distributions to allow interpolation and extrapolation. There is a clear clumping of responses at 5-year intervals; while such clumping does not substantially bias the estimates obtained below for this distribution, it was ultimately (see Section 6.4) taken into account to some extent by binning responses into a set of ranges (1, 2, 3, 4, 5, 6–10, 11–18, 19–25, 26–35, 36–45, 46–55 years, and an empty range of all larger values). This binning was found to have negligible effect on the parameter estimates. After examination of the empirical distribution, the cumulative distribution of time-eating-to-present was fitted using maximum likelihood with a parametric distribution of the form

$$F_1(t) = 1 - \alpha \exp(-\lambda_1 t) - (1-\alpha)\exp(-\lambda_2 t) \qquad (6.8)$$

where t is the time for which fish have been eaten up to the time of the survey, and α, λ_1, λ_2 are parameters. To ensure uniqueness, we also impose the condition $\lambda_1 > \lambda_2$. This two-exponential curve gives an adequate fit as indicated by the Kolmogorov-Smirnoff statistics (Knuth, 1998; K^+=0.615, p=0.46; K^-=0.830, p=0.24) computed for the fitted versus empirical curve, while a single exponential fit is clearly inadequate (K^+=1.59, p=0.006; K^-=1.04, p=0.11).[14] The empirical cumulative distribution and the fitted distributions (smooth curves) are shown in two ways in Figures 6.1 and 6.2. The second, plotting the logarithm of the complement of the cumulative fraction of eaters[15] against time, shows that the fitted curve may somewhat overestimate probabilities for large times (over about 35 years). This overestimate may only be apparent for current and future fish-eaters, however, since at periods more than about 35 years ago the river was less suitable for fishing over appreciable portions of its length.

---

[13] The five zero entries in the database have been ignored, as has the entry of 80 years attributed to a respondent aged 18-30 years. Including them leads to changes in parameter estimates that are well within the uncertainties estimated.

[14] These values are for estimates made without binning responses to the intervals described. The use of the Kolmogorov-Smirnoff statistic in this way is purely heuristic — this statistic is not designed for testing the adequacy of a parametric fit, and the data are not exactly of the form required for the test, since the precision of each data point is limited by the reported one-year intervals, and the accuracy is affected by recall biases, including the clumping at 5-year age intervals. The values given for probability (p) must thus be interpreted with extreme caution.

[15] The cumulative fraction at a given time is estimated for the purposes of plotting as the number of such eaters as reported in the survey, minus 0.5, divided by the total number responding. The subtraction of 0.5 allows plotting of all points, and approximates an unbiased location for the plotted point (Cunnane, 1978). This device is used for plotting purposes only — no such compensation is required or used in the maximum likelihood fitting.

Table 6.10 Numbers of years[a] eating fish from the Kalamazoo, and numbers of respondents.

Years	Number	Years	Number	Years	Number	Years	Number
1	45	10	27	19	3	33	2
2	34	11	4	20	16	35	5
3	28	12	2	21	4	36	1
4	17	13	1	22	2	38	2
5	22	14	1	23	3	39	1
6	8	15	11	24	1	40	3
7	7	16	2	25	5	42	1
8	4	17	2	28	1	45	1
9	5	18	3	30	11	55	1

[a] Recorded responses were integer numbers of years. Only those numbers of years with at least one respondent are shown in the table.

For the purposes of fitting distributions, these data were interpreted to indicate that the respondent's number of years eating fish up to the time of the survey lay within the range of ±0.5 year of the integer number of years specified, except that 1 year was taken to indicate any period from zero to 1.5 years. The likelihood contribution for each individual year t_i in the table is then

$$\begin{cases} n(t_i)\ln\left(F_1(t_i+0.5)-F_1(t_i-0.5)\right) & \text{for } t_i > 1 \text{ year} \\ n(t_1)\ln\left(1-F_1(t_1+0.5)\right) & \text{for } t_1 = 1 \text{ year} \end{cases} \tag{6.9}$$

where $n(t_i)$ is the number of respondents associated with year t_i.

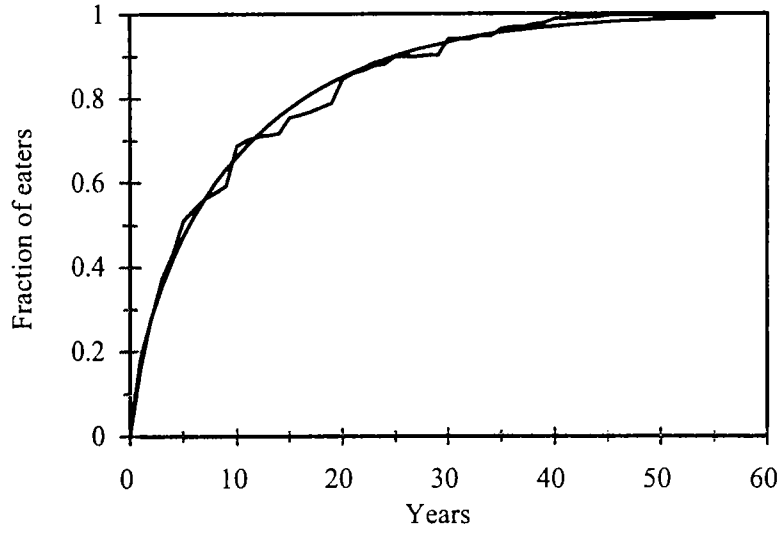

Figure 6.1 Cumulative distribution of reported years spent eating fish at the time of survey

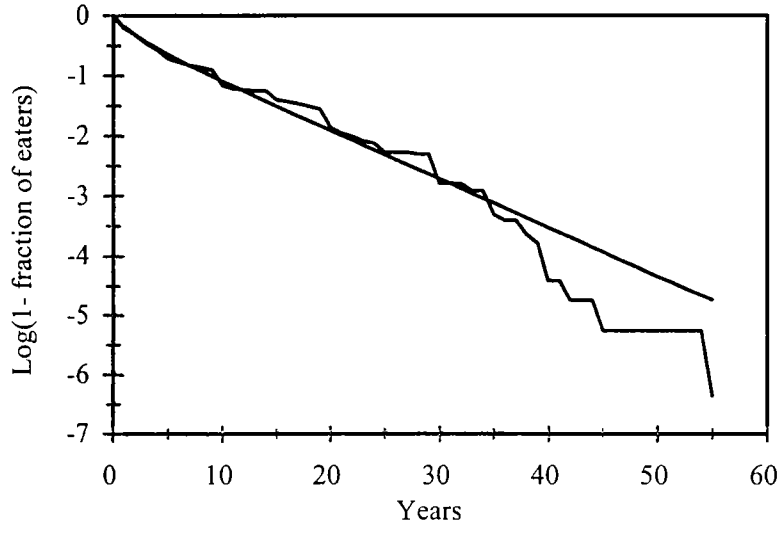

Figure 6.2 Cumulative distribution of reported time eating fish — alternative scale

The maximum likelihood estimates for the parameters are:

λ_1	0.323 per year
λ_2	0.0810 per year
α	0.269

These values were actually obtained simultaneously with the estimate for the distribution of initial ages, as explained in Section 6.4, with binning of ages to account for the observed clumping that is probably a recall artifact. (For comparison, the maximum likelihood values obtained using equation 6.9 with no binning of ages are $\lambda_1 = 0.372$ per year, $\lambda_2 = 0.0807$ per year, and $\alpha = 0.226$).

With the cumulative distribution for years to survey given as described, an estimate may be obtained for the cumulative distribution of total years in a lifetime during which fish would be eaten (Israeli and Nelson, 1992), with the assumption that these distributions do not change over time or with the age of the respondent. The cumulative distribution for total time of fish eating is then:

$$F_2(t) = 1 - \frac{\alpha\lambda_1 \exp(-\lambda_1 t) + (1-\alpha)\lambda_2 \exp(\lambda_2 t)}{\alpha\lambda_1 + (1-\alpha)\lambda_2} \tag{6.10}$$

This functional form necessarily overestimates the upper end of the distribution at very large times of fish-eating, because it has no upper bound on such times. The period spent eating fish from the Kalamazoo clearly cannot exceed a lifetime, although for any fisher it could exceed the time spent fishing — if, for example, a relative continued to bring home fish. We therefore retained this functional form for fitting the variability distributions.

The functional form as fitted to the observations has the effect of estimating quite a high probability (27.5% at the maximum likelihood values) for periods of fish-eating of less than 1.5 years in a lifetime, and 19.6% for periods less than 1 year. The similar estimate for the expected fraction to be seen in a survey at a fixed time is 18.7% less than 1.5 years, 13.1% less than 1 year, again at the maximum likelihood estimate. For comparison, the observed fraction of anglers who responded with one year as the length of time that they had been eating fish was 15.7% (45/286) who (the difference from either 18.7% or 13.1% is not statistically significant). To ensure that exposures are not underestimated, any estimate of less than one year was taken to be one year in the simulation. This adjustment is performed before estimating the number of meals per year (see Section 6.5.1).

The functional form as fitted also estimates a small probability for very long periods of fish-eating. At the maximum likelihood estimates, there is a 0.14% chance of estimating a lifetime period of fish-eating that exceeds a standard lifetime of 70 years (the corresponding estimated fraction of such fish-eaters in a survey such as the Phase I survey is 0.25%, consistent with the

observation of none). As explained in Section 6.3.2, there is no standard method of adjusting risk estimates for exposure periods exceeding a standard lifetime. Here, we truncate the potential exposure period at the standard lifetime of 70 years — if the estimated exposure period exceeds a standard lifetime, it is set to the standard lifetime.

The uncertainty in the variability distribution may be estimated from the maximum likelihood procedure, and is summed up by the uncertainties in the parameter estimates λ_1, λ_2, and α. The joint uncertainties in these were estimated using the inverse of the information matrix (the matrix of second derivatives of the loglikelihood function, evaluated at the maximum likelihood values) as an estimator of the variance-covariance matrix, and used to construct a multinormal uncertainty distribution. Since this procedure was carried out simultaneously with the estimate of the initial age of fish-eating, the variance-covariance matrix so constructed is described in Section 6.4.

This approach accounts for the uncertainty in the parameter estimates, but does not completely incorporate some further uncertainties:

- The functional form for the distributions of time-spent-eating. Most of the uncertainty in functional form should be incorporated in the current uncertainty estimates, because the fit to the empirical data is very good.
- The assumption of constancy in time and age. It is assumed that the distribution of time-spent-eating is independent of calendar time, and of initial age at which fish-eating starts. There is some indication in the data that the approach taken overestimates the number of both long-term-eaters and of short-term, young fish eaters.

6.3.2 Effective additional exposure period for anglers

PCBs accumulate in the body, and are metabolized or excreted at a relatively slow rate. Less-than-lifetime exposure to PCBs thus results in a body burden that remains after the time of exposure, and which may contribute to cancer risk. The standard estimate of cancer risk is based on a standard lifetime (70 years) of continuous exposure to PCBs, so corresponds to the average body burden during that standard lifetime. To estimate the effective additional exposure resulting from the remaining body burden after less-than-lifetime exposure requires comparing average body-burdens over the standard lifetime of 70 years; it may not suffice to simply estimate the average lifetime exposure by averaging intake over the standard lifetime.

Because individual PCB congeners are removed from the body at different rates, the total accumulation of PCBs in the body is dependent on the mix of congeners. The non-metabolic excretion rates for all congeners, and the metabolic rates for 146 PCB congener combinations, corresponding to the peaks from DB-1 capillary gas chromatography columns (by Northeast Analytical, Inc., Schenectady, NY), have been determined or estimated by Brown (1994). The 146 peaks include at least 185 of the 209 PCB congeners (24 congeners were not assigned to any

particular peak, but 13 peaks contained unidentified congeners). We assigned a total removal rate to each congener as the sum of the non-metabolic rate and the normal metabolic rate[16] for the DB-1 peak containing that congener (see the spreadsheet PCB_congener_data.wb3, Appendix B.13). Where only lower bounds or approximations on the metabolic rate constant were provided, those lower bounds or approximations were used. For the 24 congeners not assigned to a peak, this results in an underestimate of the removal rate (since we set the metabolic rate to zero) — but 19 of these congeners did not occur in detectable quantity in published measurements of Aroclors 1016, 1242, 1248, 1254 or 1260 (Frame *et al.*, 1996; ATSDR, 2000), and for the 6 that were detected the largest mass fraction was 0.23%. Indeed, Frame *et al.* (1996) report that 52 congeners were not detected above 0.01% by weight in any of the 17 Aroclor mixes they analyzed.[17] The four 2,3,5,6 tetrachlorophenyl-containing congeners (one ring contained 2,3,5,6 chlorination at least, IUPAC numbers 178, 193, 202, and 208) that increased relative to the others (presumably due to dechlorination of more chlorinated congeners) were assigned a metabolic removal rate of zero. In the published measurements, these four congeners contributed less than 0.1% mass fraction to Aroclor 1254, 1.82% to Aroclor 1260, and undetectable quantities to other Aroclors..

The sum of the congener metabolic and non-metabolic removal rates can be used to calculate the accumulated amount of PCBs in the body over time due to the intake of any mixture of congeners. The body burden of a particular PCB congener after a dosing period t is given by

$$B_i(t) = \frac{r_i}{k_i}\left(1 - \exp\left(-k_i t\right)\right) \tag{6.11}$$

where the terms and their dimensions are

B_i	Body burden of congener i (dimensionless),
r_i	Dose rate of congener i (T^{-1}),
k_i	Total metabolic and non-metabolic clearance rate for congener i (T^{-1}), and
t	Accumulation time (T),

so the total body burden B is

$$B(t) = \sum_{i=1}^{209} \frac{r_i}{k_i}\left(1 - \exp\left(-k_i t\right)\right) \tag{6.12}$$

---

[16] Brown (1994) provides estimated metabolic rates for normal and chloracnegenic persons, the latter having been exposed to sufficiently high doses of PCBs as to induce higher metabolic rates. The normal rates are thus smaller, and result in estimates of higher body burdens.

[17] It is possible that Brown mis-identified some of the congeners that occur in small quantities, or that do not occur in Aroclors, in some of the chromatogram peaks. The effect of such mis-identifications would be negligible in this analysis.

These equations have been derived assuming a constant body mass, expressing the body burden as a fraction of total body mass, and the dose rate as a fraction of total body mass per unit time (*e.g.* corresponding to units such as mg/kg-day). This approximation is sufficiently accurate for the purposes of this assessment (it results in a slight overestimate of body burden, since it omits the dilution due to growing body mass).

The time integral of the body burden (the "area under the curve" of the body burden) for congener i after time t at a dose rate r_i is then given by

$$h_i(t) = \frac{r_i}{k_i}\left(t - \left(1 - \exp\left(-k_i t\right)\right)/k_i\right) \tag{6.13}$$

so that the integral of the total PCB body burden is (summing over the 209 PCB congeners)

$$h(t) = \sum_{i=1}^{209} \frac{r_i}{k_i}\left(t - \left(1 - \exp\left(-k_i t\right)\right)/k_i\right) \tag{6.14}$$

assuming no interactions between congeners in adsorption, distribution, metabolism or excretion. This non-interaction assumption was the basis for the measurement of the rates, and appears to be adequate except for the four previously mentioned congeners that apparently increased relative to the others due to de-chlorination of higher-chlorinated congeners. At the low dose rates of PCBs in the Kalamazoo fish-eating population (low compared with the dose rates in the situations that Brown, 1994, used to estimate the rates), no such interactions are expected.

The lifetime average body burden, H, due to dosing from age t_1 to age t_2 during a standard lifetime of T may then be expressed as:

$$H\left(t_1, t_2\right) = \frac{h\left(T - t_1\right) - h\left(T - t_2\right)}{T} \tag{6.15}$$

The average body burden from a lifetime exposure (age 0 through T) can be found by setting t_1 = 0 and $t_2 = T$.

An effective exposure duration can be found by taking the ratio of the lifetime average burden from a less-than-lifetime exposure, $H(t_1, t_2)$, with the burden from a standard lifetime exposure, $H(0, T)$, and multiplying this value by the standard lifetime of T (70 years). The change necessary to give the correct effective years of exposure, the effective additional exposure period t_e, is then the difference between the calculated effective exposure duration and the actual period of exposure:

$$t_e = \frac{H\left(t_1, t_2\right)}{H(0, T)} \cdot T - \left(t_2 - t_1\right) \tag{6.16}$$

This value has been calculated for less-than-lifetime exposures. The congener mixture was chosen to correspond to 75% smallmouth bass + 25% carp (using the average of all fish fillet data collected in ABSAs 3 through 9, treating non-detects as ½ the detection limit). While the

effective additional exposure period does depend on the congener mixture, it does not vary strongly with small changes. The modification to the exposure period is generally less than 10%, so that the calculated changes in effective exposure period with this fixed congener mix are adequate for the Aroclor mixes likely to arise from any combination of fish from the Kalamazoo in the diet. The Aroclor fractions used are:

Table 6.11 Aroclor composition used to estimate change in effective exposure period.

Aroclor	1221	1232	1016	1242	1248	1254	1260
Average fraction 75% bass + 25% carp	0.0%	0.0%	9.8%	11.1%	17.9%	52.9%	8.3%

This Aroclor composition was used in conjunction with the averages of the Aroclor congener compositions obtained by Frame, *et al.* (1996; see also ATSDR, 2000) to obtain an estimate of the congener composition of the Aroclor mix represented by Table 6.11. The calculation is performed in the spreadsheet PCB_congener_data.wb3 (Appendix B.13).

Figure 6.3 shows the calculated effective additional exposure duration as a function of the actual exposure duration and initial age. As can be seen, for some combinations of initial age and exposure duration, the effective additional exposure duration is negative. This is a necessary consequence of the assumption that the cancer risk is proportional to lifetime average body burden. For example, consider the effect of two exposures at equal dose rates, one from birth to age 35, the other from age 35 to 70. A person (call him Alex) suffering both exposures would be exposed for a full lifetime. A person (call him Bill) exposed for the first half of a lifetime would have a continuing body burden after cessation of exposure, so that his average body burden would be higher than ½ the lifetime average of Alex — on Figure 6.3 the effective additional exposure duration is positive (about +7.4 years, so his lifetime average body burden is about (35+7.4)/70≈0.61 of Alex's). On the other hand, a person (call him Carl) exposed for the second half of a lifetime must then have an average body burden less than half of Alex — because the lifetime average body burdens of Bill and Carl must add up to give the same as the lifetime average body burden for Alex. On Figure 6.3, Carl would have a negative effective additional exposure duration (about -7.4 years, so his lifetime average body burden would be about (35–7.4)/70≈0.39 of Alex's).

The curves of Figure 6.3 can be adequately represented by using the following empirical approximation for the function $h(t)$. This approximation was computed for the congener fractions given in Table 6.11. The parameter values were chosen so that h is normalized to 70 years at $t = 70$ years:

$$h(t) \approx f_m \left(t - \frac{1 - \exp\left(-f_0\left[1 - \exp\left(-f_1 t\right)\right] - f_2\left[1 - \exp\left(-f_3 t\right)\right] - f_4 t\right)}{f_0 f_1 + f_2 f_3 + f_4} \right) \tag{6.17}$$

where the values and dimensions of the parameters are:

f_m	=	1.48273	dimensionless
f_0	=	0.001623	dimensionless
f_1	=	1.38917	per year (T^{-1})
f2	=	0.09715	dimensionless
f3	=	0.10141	per year (T^{-1})
f_4	=	0.02474	per year (T^{-1})

With this approximation for *h*, the maximum error in estimates for excess time, using equation 6.15, is less than 0.02 years.

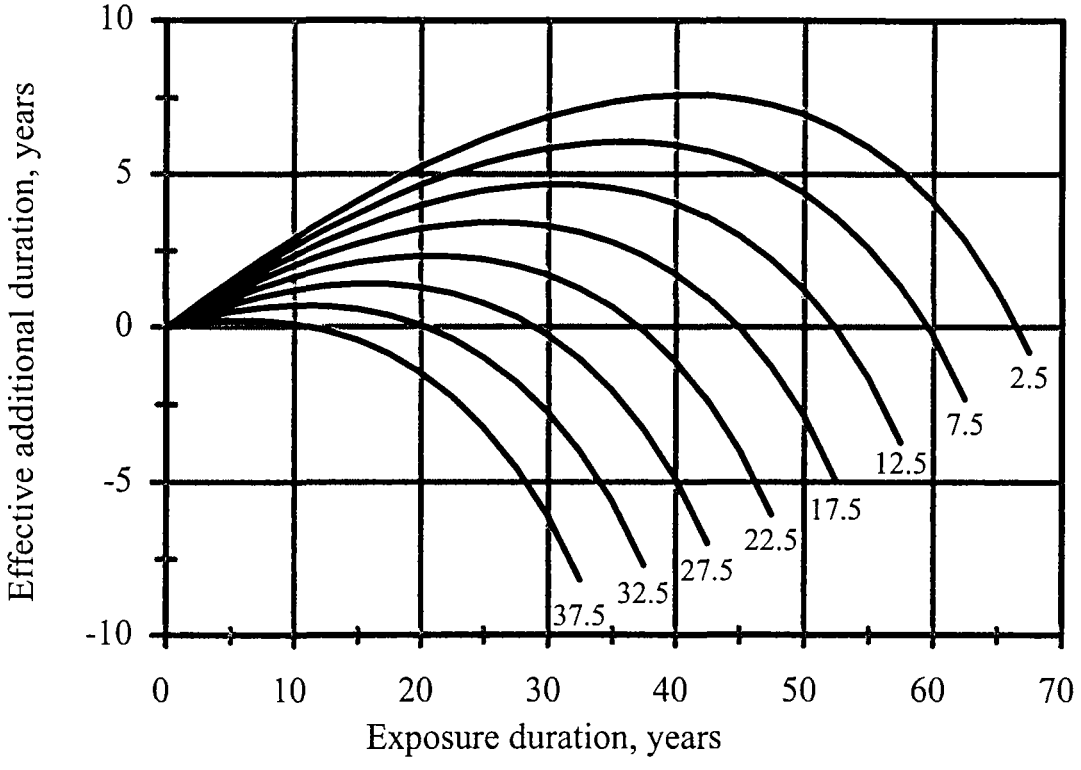

Figure 6.3 Effective additional exposure duration. Each line is labeled with the initial age of exposure.

The curves shown in Figure 6.3 indicate what happens for exposures that terminate at or before a standard lifetime. For exposures that continue beyond a standard lifetime there is no standard approach to estimating lifetime risk. For simplicity, and to avoid underestimating risks, we treat such cases as though the exposure started earlier in life and terminated at a standard lifetime (70 years).

The uncertainty in the estimated effective additional exposure duration cannot be obtained from any experimental data, since no such data exist. Indeed, the concept of averaging dose rates at different periods during a lifetime in order to estimate an average lifetime dose is based on an untested hypothesis that the average obtained is a dose rate metric that is proportional to risk. Thus modifications of the procedure are also untested hypotheses. Rather than explicitly incorporating any additional uncertainty, the uncertainty of this procedure is subsumed in that due to the use of standard cancer potency factors (incorporating the assumptions that low dose rates cause cancer, at rates that may be predicted from a linear extrapolation from the ED_{10} of

animal studies), and so applicable to any risk assessment of this nature. To evaluate the potential size of any effect of incorporating this term, the sensitivity analysis (Section 6.11.1) includes an evaluation of the effect of omitting the term entirely — setting the effective exposure duration to the actual exposure duration.

6.4 *Initial age of fish eating*

The distribution of initial ages at which Kalamazoo anglers start eating fish can be estimated from the responses obtained in Phase I of the Kalamazoo River Angler Survey (MiCPHA, 2000a,b), by extending the analysis described in Section 6.3.1. Fish eaters, in addition to providing information on the length of time for which they had been eating fish, indicated their age, at least in broad ranges. This information on age, when combined with the information on period of eating fishes, enables an estimate to be made of the distribution of initial ages for starting to eat fish. Table 6.12 shows the combined period versus age information obtained in the Phase I survey (MiCPHA, 2000a,b).

Under the conditions described in Section 6.3.1, the expected number of respondents in an age range from s_1 to s_2 who had eaten fish for a length of time somewhere in the range T_1 and T_2 would be:

$$rN\int_{s_1}^{s_2} ds\int_{T_1}^{T_2} dT\, q(s-T)\big(1-F_2(T)\big) \tag{6.18}$$

where the terms are:

- r the rate of entry to the fish-eating population (T^{-1})
- N the total current population (dimensionless) of fish-eaters (the number of persons who still eat fish from ABSAs 3 through 9 of the Kalamazoo at any one time),
- $q(t)$ the probability density function, evaluated at initial age t, for the distribution of initial ages of fish-eaters (T^{-1}), and
- $F_2(T)$ the cumulative probability (dimensionless) for eating fish for a length of time T.

Integrating over all initial ages and all lengths of time eating, we see that

$$r=\left[\alpha\lambda_1+(1-\alpha)\lambda_2\right] \tag{6.19}$$

using the functional form for F_2 given in Section 6.3.1, equation 6.10. With the parameter values given in Section 6.3.1, the rate of entry to the fish-eating population is about 15% per year of the current total number eating fish. This estimate is used in Section 6.8.

The information on age ranges is limited in the Phase I survey (below 18, 18–30, 31–45, 46–60, over 60), but the risk assessment is insensitive to the distribution of initial ages — the estimate of initial age is used only in estimating the small correction to the duration of eating fish in order to account for PCB accumulation (Section 6.3.2). The integral for the expected number of

respondents in given age ranges for given periods of fish eating (equation 6.18) was evaluated numerically (see spreadsheet Age_structure.wb3, Appendix B.14), and a likelihood function constructed by noting that the numbers obtained in the Phase I survey in each age-range and period of eating fish formed a multinomial distribution, with expected values given by the integral.

The information on ages and time spent eating fish is not necessarily entirely accurate, as already noted in Section 6.3.1. There is clearly clumping of the estimates for time spent eating fish at five year intervals; and self-reported information on age may also be inaccurate. To take account of the clumping of estimates at 5 year intervals, the responses on the length of time spent eating fish were binned to the values 1, 2, 3, 4, 5, 6–10, 11–18, 19–25, 26–35, 36–45, 46–55, and implicitly >55 years, with no responses in this last bin. This binning has a very minor effect on the estimates for the distribution of time spent eating fish, as mentioned in Section 6.3.1.

Table 6.12 Numbers of years[a] eating fish from the Kalamazoo, and numbers of respondents, by age groups <18,18–30,31–45,46–60,>60.

Years	Number	Years	Number	Years	Number
1	2,20,15,8,0	13	1,0,0,0,0	25	0,0,2,2,1
2	11,6,6,9,2	14	0,0,1,0,0	28	0,1,0,0,0
3	4,10,8,2,4	15	2,4,2,2,1	30	0,1,8,2,0
4	2,4,4,3,4	16	0,1,0,1,0	33	0,0,2,0,0
5	1,5,10,3,3	17	2,0,0,0,0	35	0,0,1,2,2
6	0,4,4,0,0	18	0,1,0,0,2	36	0,0,0,1,0
7	1,0,4,1,1	19	0,1,1,1,0	38	0,0,0,2,0
8	0,2,1,1,0	20	0,3,6,6,1	39	0,0,0,1,0
9	1,0,1,2,1	21	0,1,2,1,0	40	0,0,1,1,1
10	3,7,8,6,3	22	0,0,0,0,2	42	0,0,1,0,0
11	0,1,3,0,0	23	0,0,1,0,2	45	0,0,0,0,1
12	0,1,1,0,0	24	0,1,0,0,0	55	0,0,0,1,0

[a] Recorded responses were integer numbers of years. Only those numbers of years with at least one respondent are shown in the table.

The responses of several respondents' for age and time spent eating fish that they caught are incompatible — they imply that those respondents caught and ate fish at ages 0 through 5. While respondents certainly may have eaten fish from the Kalamazoo at such ages, it is highly unlikely that they regularly ate self-caught fish at those ages. In order to obtain plausible estimates of the distribution of initial ages for eating of self-caught fish, and since the risk estimates ultimately

computed depend only to a small extent on this distribution, the age group for nine respondents was modified. For these nine respondents, the reported combination of age group (below 18, 18–30, 31–45, 46–60, over 60) and time spent eating self-caught fish implied a maximum initial age of zero through 5 years (for these respondents, the maximum possible initial ages would have been, [1,1,3,3], [0,2], [3,5], [5] years, where the nine values have been grouped by the reported age ranges and there was no such respondent in the oldest age range). For the purposes of the following analysis, the age range for these nine respondents has been moved to the next higher value. This modification has no effect on the estimates of time spent eating fish, but changes the estimates of the initial-age distribution.

Non-parametric piecewise-linear estimates of the initial age distribution showed that failure to modify the age-range data resulted in a very large estimated probability for an initial age less than 5, followed by a strong dip at age around 10. With the modification, the non-parametric estimate of initial age distribution increased relatively smoothly from zero through 10, was roughly constant through age 20 to 30, then declined to zero at about age 60. A gamma distribution (Figure 6.4) was found to adequately represent this structure:

$$q = \frac{t^{\kappa} \exp\left(-\mu\left(t/t_0\right)\right)}{\left(\mu/t_0\right)^{-(\kappa+1)} \Gamma(\kappa+1)} \tag{6.20}$$

where

- q is the probability per unit age (T^{-1}) for an initial age t,
- t is initial age (T),
- t_0 = 20 years is a convenient normalization (T),
- κ is a dimensionless parameter with maximum likelihood estimate 1.477, and
- μ is a dimensionless parameter with maximum likelihood estimate 1.654.

The uncertainty in this distribution of initial age of fish-eating is obtained from the likelihood estimation technique used to obtain the parameter estimates. The uncertainties in the parameter estimates were obtained by using the inverse of the information matrix (the matrix of second derivatives of the loglikelihood function evaluated at the maximum likelihood values) as an estimator of the variance-covariance matrix, and constructing a multinormal distribution based on this variance-covariance matrix. Since the estimates were obtained simultaneously with the parameter estimates for the duration of fish-eating, the five-parameter variance-covariance matrix for λ_1, λ_2, α, κ, and μ was obtained, and the uncertainty distribution estimated by a 5-dimensional multinormal distribution (Devroye, 1986).[18]

[18] The "multinormal" distribution has a density that is proportional to the exponential of minus a quadratic form in the vector of variates. This distinguishes it from the many other "multivariate normal" distributions — multivariate distributions with normal marginal distributions.

The estimated lower diagonal half of the (symmetric) variance-covariance matrix was (when all ages and times are expressed in years):[19]

	λ_1	λ_2	α	κ	μ
λ_1	7.816e-03				
λ_2	3.527e-04	8.463e-05			
α	-5.900e-03	-6.919e-04	1.038e-02		
κ	-9.572e-06	5.316e-05	-1.338e-04	7.675e-02	
μ	-2.573e-07	3.500e-05	-1.207e-04	5.647e-02	4.674e-02

which gives the following values for the standard deviations (along the leading diagonal) and the Pearson product-moment correlation coefficients (below the diagonal):

	λ_1	λ_2	α	κ	μ
λ_1	0.0884082				
λ_2	0.4336497	0.0091995			
α	-0.65504	-0.738185	0.1018829		
κ	-0.000391	0.0208582	-0.004739	0.2770312	
μ	-0.000013	0.0175985	-0.005481	0.9428286	0.2161951

The groups λ_1, λ_2, α, and κ, μ are practically independent of one another — the small correlations shown may be due to the numerical approximations used to make these estimates. The parameters κ and μ are very highly correlated ($\rho = 0.94$). With these standard deviations, and with the multinormal approximation for the uncertainty distribution, samples can occasionally (approximately 0.7% of the time) be drawn with parameter values that are out of bounds (in

[19] This matrix of values uses the computer equivalent of scientific notation for numbers. The value d.dddesdd has to be interpreted as d.ddd × 10^{sdd} where d stands for a digit, and s for a + or – sign. For example, 8.038e-07 is the same as 8.038×10^{-7}.

particular with $\alpha \leq 0$ or $\lambda_1 < \lambda_2$)[20]. The multinormal distribution was censored to ensure that all five parameters were positive, $\lambda_1 \geq \lambda_2$, and $\alpha \leq 1$.

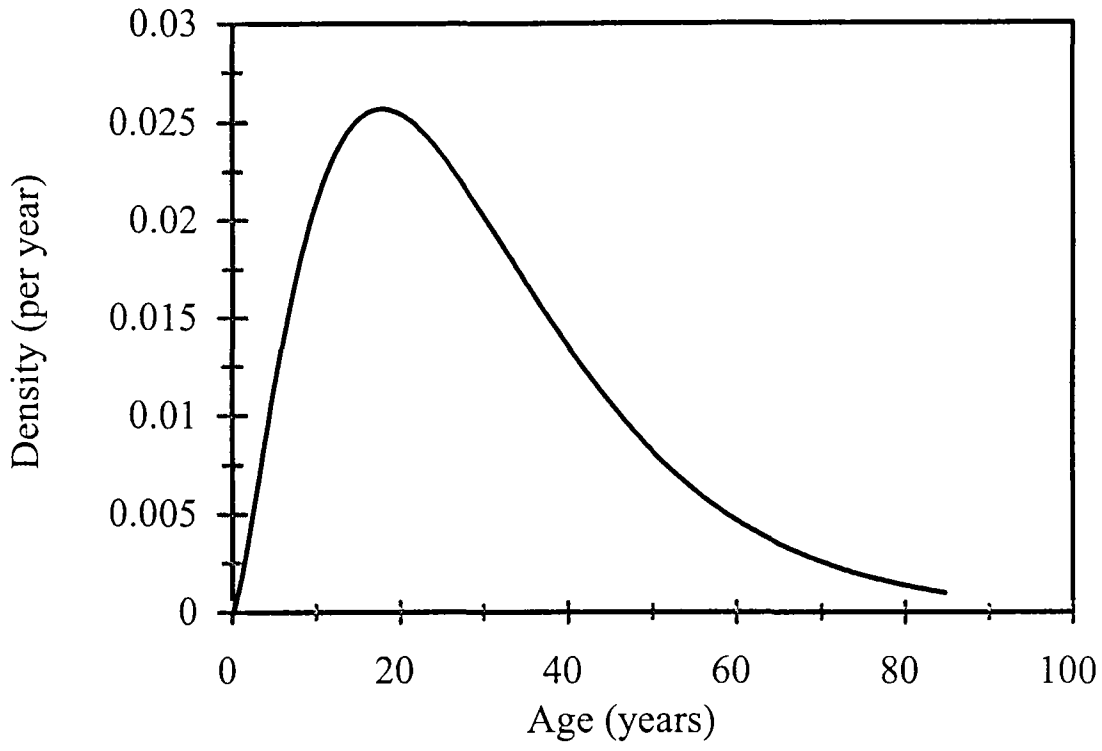

Figure 6.4 Density function for the distribution of initial ages for eating fish

[20] At $\lambda_1 = \lambda_2$ the identities of these two parameters becomes ambiguous, and this condition is equivalent to $\alpha = 0$. Using the likelihood function directly, this occurs with probability only 0.002, so the multinormal approximation is somewhat overestimating the probability for more extreme values, and so slightly overestimating the uncertainties involved.

6.5 Kalamazoo River fish consumption

6.5.1 Meals eaten per year

The Kalamazoo River Angler Survey (MiCPHA, 2000a,b) was used to estimate fish consumption (see spreadsheet Meals.wb3, Appendix B.15). The responses to the following survey questions have been used (these are the main question headers — there are subsidiary parts of the questions not quoted here, see Appendix B.3):

5. Do you or other members of your household eat the fish or snapping turtles you catch from the Kalamazoo River or Portage Creek?
6. What species of fish do your or others in your household eat from the Kalamazoo River or Portage Creek?
8 How often are the fish you catch in the Kalamazoo River or Portage Creek eaten by you or others in your household?

Of the 939 interviewees included in the available database, 352 either responded that they consume fish or turtles which they catch, or provided a length of time eating fish. Of those not providing a length of time eating, 399 specified that they ate neither fish nor turtles, 2 only said that they did not eat fish, 18 only said that they did not eat turtles, and 168 did not provide any response on whether they ate fish or turtles. Four anglers who did not indicate eating fish or turtles later responded positively to questions about consumption of individual types of fish that they caught. In addition, 294 said that members of their household eat fish or turtles from the river or creek. The total number of other household members reported as eating the fish from the river and creek was 807.

The responses to question 6 included the "approximate number of meals per year" by fish species and apparently, although not indicated on the questionnaire, sometimes also or in addition a total number of meals per year (data provided in the response information, MiCPHA, 2000b). The total number of meals ranged from 1 to 365 per year. Question 8 obtained broad estimates in ranges of "less than once per month", "approximately once per month," "approximately once per week" and "more than once per week." Both questions requested information about both the angler and the household, although the number of meals per year was not so separated in question 6 — we evaluate here the responses for the angler. It is assumed that the minimum number of meals per year is one — any fewer, and the angler could scarcely be considered a regular fish eater.

Of those questioned, 294 provided some indication of the number of meals eaten per year. We took the best estimate of the total number of meals per year to be the total number provided (if there was one), otherwise the sum of the numbers of meals by fish species (this was usually the same as the total number, if both were provided), otherwise within ranges 0 through 12, 12 through 26, 26 through 52, or 52 to 1000 (the last value chosen sufficiently large that it makes no material difference to the analysis), for the four responses in question 8.

Of the 294 anglers with some indication of the number of meals per year, 265 also provided an estimate of the length of time for which they had been eating fish at the time of the survey. There is a small correlation between the reported length of time and the average number of meals per year — using point estimates for the number of meals per year (actual number where given, or 6, 12, 52, and 100 for the ranges) the Pearson correlation coefficient is about 0.22. We therefore evaluated the number of meals per year conditioned on the reported length of time eating fish, on the assumption that anglers remain approximately consistent in their habits throughout the period during which they eat the fish they catch.[21] The distribution of number of meals per year for the 29 fish-eating anglers who did not report the length of time for which they had been eating fish did not appear different from the 265 who did report that information (no formal statistical tests were conducted). The potential bias arising from the differential probability of inclusion in the Phase I survey due to different frequency of fishing has been ignored, because the probability differences are not too large (except perhaps for very infrequent fishers, see Section 6.7), and the correlation between frequency of fishing and number of meals per year appears to be relatively small (spreadsheet Meals.wb3, Appendix B.15, shows that the Pearson correlation coefficient between meals/year and the reported number of times fishing in the last calendar year is about 0.07).

To estimate the distribution of meals per year eaten, the empirical distributions were plotted for approximate quartiles of the years eating fish.[22] These corresponded to groups of anglers eating fish for 1 or 2 years (67 anglers), for 3 through 5 years (64 anglers), for 6 through 16 years (67 anglers), and for 17 through 55 years (66 anglers) — Figure 6.5. It was found that the duration-eating-fish-conditioned distributions could be adequately fit by censored lognormal distributions of the form:

$$P_t(n) = \frac{\Phi\left(\left(\ln(n) - \mu(t)\right)/\sigma(t)\right) - \Phi\left(-\mu(t)/\sigma(t)\right)}{1 - \Phi\left(-\mu(t)/\sigma(t)\right)} \quad \text{for } n \geq 1 \qquad (6.21)$$

where

[21] There is a slight mismatch between the analysis performed here and the way it is used in the Monte Carlo simulation, in that the correlation obtained here is with the period of time spent eating fish up to the time of the survey, whereas the simulation uses the period of time spent eating fish during a lifetime. This mismatch appears to be unavoidable without excessive amounts of computation, since for individual anglers the latter information cannot be obtained in a survey. Since the correlation is relatively small, the effect of the mismatch is also small. The mismatch is in a direction that overestimates the amount of fish eaten, since the correlation is positive — the longer the period spent eating fish, the larger the estimated number of meals per year; and the lifetime period spent eating fish necessarily equals or exceeds the period at the time of any survey.

[22] As in Section 6.3.1, the reported duration of 80 years was omitted, as were the reported durations of 0 years. The number of anglers reported in the text, 265, included the angler indicating 80 years duration, but omitted those reporting 0 years.

$P_t(n)$ is the cumulative probability (dimensionless) for n meals per year for an angler with eating duration t,
$\mu(t)$ is a mean value (dimensionless) that increases with t,
$\sigma(t)$ is a standard deviation (dimensionless) that decreases with t, and
Φ is the standard cumulative normal function.

Maximum likelihood estimates were obtained for μ and σ as linear functions of the logarithm of t. The likelihood was based on the functional form for $P_t(n)$, using point estimates for n where responses to question 6 were available, and range estimates where only responses to question 8 were available. The empirical fits for μ and σ were of the form :

$$\begin{aligned} \mu(t) &= \mu_a + \mu_b \ln(t/t_0) \quad \text{for} \quad t \geq t_0 \\ \sigma(t) &= \sigma_a + \sigma_b \ln(t/t_0) \quad \text{for} \quad t \geq t_0 \end{aligned} \tag{6.22}$$

where $t_0 = 1$ year is included to standardize units of time, and the maximum likelihood estimates for the parameters are

μ_a	=	0.875	σ_a	=	1.418
μ_b	=	0.675	σ_b	=	–0.0652

The smooth curves shown in Figure 6.5 correspond to the distributions obtained for value of t of 1.5, 4, 10, and 22 years — these are approximately the medians for the durations of the four quartiles plotted, but the curves are plotted principally to show the general agreement between the fits and the shape of the empirical observations (stepped curves in Figure 6.5). The minimum estimated number of meals per year is 1, corresponding the Survey questionnaire (and this response was quite common). To slightly increase the realism of the simulations, the number of meals per year was rounded up slightly so that the number of meals in a lifetime was an integer. Similarly to increase realism, the number of meals per year was limited to 1095 (3 × 365), although, as can be seen from the distributions, the effect of this is negligible.

The uncertainty in the distributional estimates for number of meals per year was estimated by estimating the variance-covariance matrix for μ_a, μ_b, σ_a, and σ_b from the inverse of the information matrix (the matrix of second derivatives of the loglikelihood function evaluated at the maximum likelihood estimates), and using that variance-covariance matrix to construct a multinormal uncertainty distribution. The lower diagonal of the (symmetric) variance-covariance matrix obtained was:

	μ_a	μ_b	σ_a	σ_b
μ_a	0.05694			
μ_b	-0.02031	0.00874		
σ_a	-0.01614	0.00502	0.01939	
σ_b	0.00392	-0.00127	-0.00655	0.00324

which gives the following values for the standard deviations (along the leading diagonal) and the Pearson product-moment correlation coefficients (below the diagonal):

	μ_a	μ_b	σ_a	σ_b
μ_a	0.2386			
μ_b	-0.0080	0.0935		
σ_a	-0.4857	0.3858	0.1393	
σ_b	0.2888	-0.2394	-0.8262	0.0570

The maximum likelihood estimate for σ_b is negative, and its standard deviation is about the same size so that samples can be more than twice as negative. With low probability, the estimated standard deviation for the distribution of the logarithm of the number of meals per year could become relatively small, or even negative, for large durations of eating fish. The entire multinormal uncertainty distribution was therefore censored to restrict $\sigma(t)$ to be greater than 0.5 at a fish-eating duration of one standard lifetime (that is, $\sigma(70) > 0.5$).

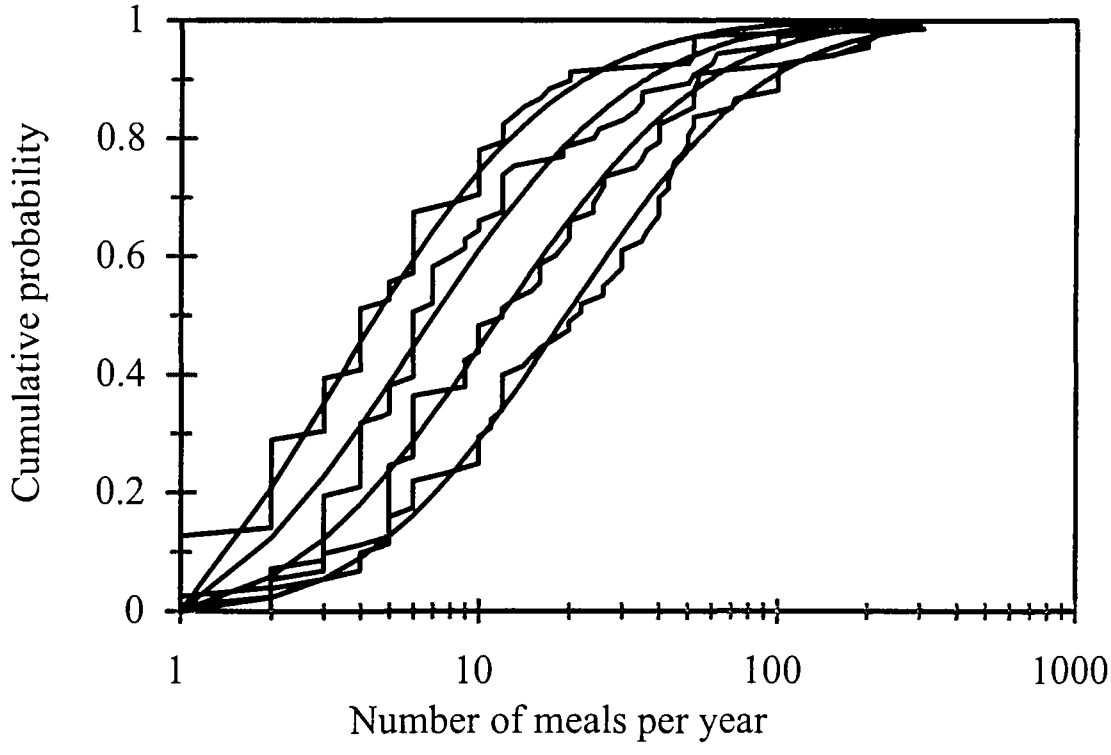

Figure 6.5 Distributions of number of Kalamazoo fish per year, conditioned on quartiles of duration of eating fish. Longest duration on the right, shortest on the left.

6.5.2 Types of fish consumed

Question 6 of the Kalamazoo River Angler Survey (MiCPHA, 2000a,b) requested information on the number of meals per year consumed of walleye, suckers, carp, bass (small mouth and large mouth), pike, panfish (perch, crappie, bluegill, sunfish), catfish, bullheads, snapping turtles, and other. Full information was provided by 237 respondents, allowing estimation of the fraction of meals of each species eaten by each angler (see spreadsheet Meals.wb3, Appendix B.15). No information on concentrations of PCBs in bullheads in particular is available, so these were subsumed into the catfish (bullheads are in the same family).

Figure 6.6 shows the average fraction of each fish species consumed by anglers who eat fish, separately by approximate quartiles of meals per year consumed (1–3, 4–7, 8–20, and 22–365 meals per year in the 1st through 4th approximate quartiles respectively, with 57, 58, 61, and 61

respondents respectively). These averages are computed on a "per angler" basis — they are the average over anglers in each quartile of the average for each angler. The figure shows that there is no great variation in average meal fraction with the number of meals per year eaten; however, the angler-to-angler variability may be substantial — particularly for small numbers of meals per year. The variability of these average meal fractions was represented by using the empirical data directly, split by quartiles of the number of meals per year. Thus, in the Monte Carlo sampling, each simulated angler (each of the innermost iterations described in Section 6.1) was assigned a number of meals per year from the distribution described in Section 6.5.1. This allowed selection of the quartile of meals per year (using cutoffs of <3.5, $3.5 < 7.5$, $7.5 < 21$, ≥ 21 meals per year), and one of the angler records in that quartile was randomly selected with equal probability within the quartile to represent the meal fractions of each type of fish for the simulated angler. Since no information is available on concentrations of PCBs in "Other" fish, any fraction of meals assigned to the "Other" category was reassigned to all the named fish in proportion to their average fraction within the quartile of the simulated angler.

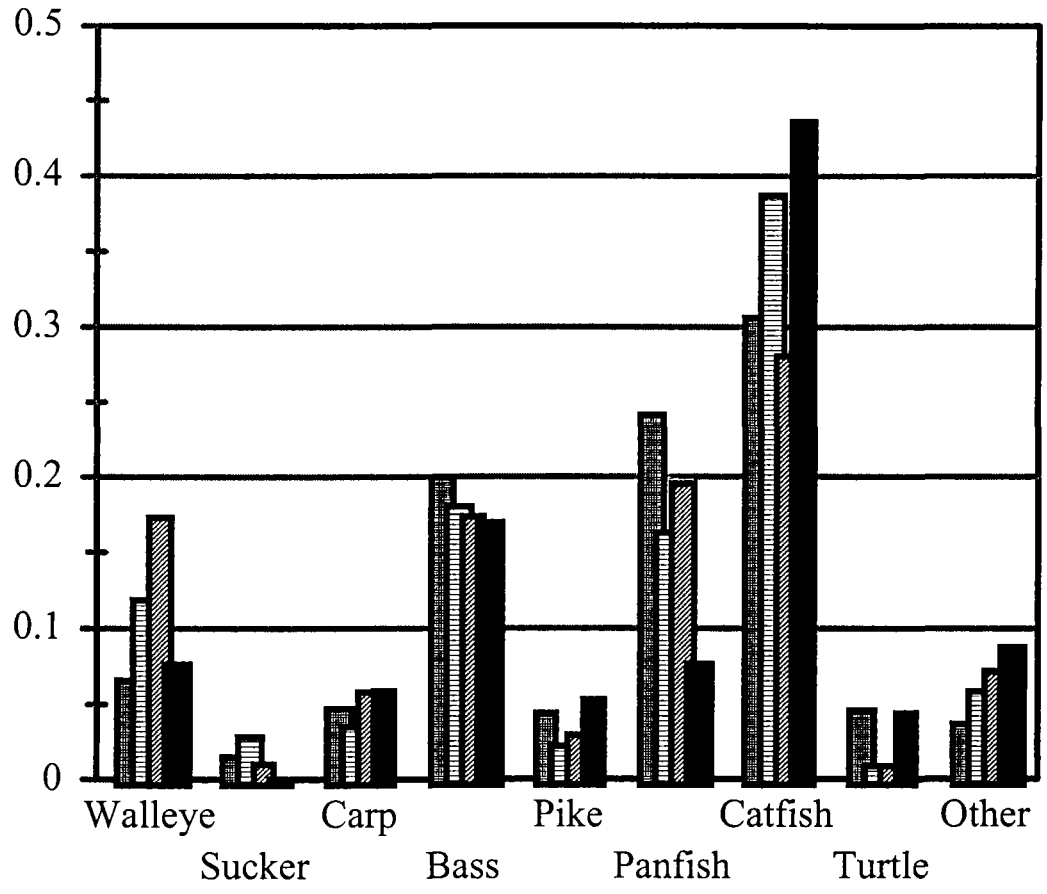

Figure 6.6 Average fraction of meals for each species of fish, by approximate quartile of meals per year. (Fourth through first quartile, left to right)

6.5.3 Meal size

There are no direct measurements of meal size in the Kalamazoo River angler population, nor questions directed at obtaining estimates of meal sizes included in Phase I of the Kalamazoo River Angler Survey (MiCPHA, 2000a,b). Phase II of this survey did request information on fish size portions, but the representativeness of that part of the survey is questionable.

Another source of information relevant to fishing in the Kalamazoo River is a telephone interview survey (Atkin, 1994; see spreadsheet Atkin_survey.wb3 and associated codebook Fish_Codebook.doc, Appendix B.17) of 690 anglers residing near the Kalamazoo River basin (out of 981 who were contacted).[23] This survey obtained representative information on fish meal sizes. In that survey, anglers were asked how many meals of different types of self-caught fish they had eaten in the previous two weeks, and were asked to estimate their size as a "small portion... say, four or five ounces, or a large amount greater than ten ounces, or in between". Atkin considered that the small portion could be adequately represented by 4 ounces, the in-between portion by 8 ounces, and the large portion by 10 ounces. Then for the 177 anglers who had eaten self-caught fish and provided information on both number of meals and meal sizes, the distribution of average meal size is given in Table 6.13 (see also the spreadsheet Meals.wb3, Appendix B.15).

Table 6.13 Average meal size by number of anglers in the Atkin (1994) survey.

Average meal size (oz) [a]	Number of anglers
12	57
10.7	1
10.4	1
10	2
9	1
8	82
6.7	5
4	28

[a] Meal sizes intermediate between 8, 10, and 12 oz. occur because some anglers reported different meal sizes for different types of fish; these averages are weighted by the relative amounts of fish eaten.

[23] The data file for this survey contains records for just 689 anglers.

Atkin (1994) did not explicitly specify whether the portion size was considered to be before or after cooking; we assume that it is before cooking, and use this distribution of values (Table 6.13) in the Monte Carlo program to estimate variability between individuals in long-term average meal size. Examination of Atkin's data indicates that there was not much difference in the distribution of meal size by type of fish, so no differentiation was made.

The average serving size implied by the values from the Atkin (1994) survey is 8.7 ounces. In Phase II of the Kalamazoo River Angler Survey (MiCPHA, 2000a,b), 80 respondents gave estimates of serving size based on a 4-ounce model portion of fish on a 9-inch dinner plate. The responses were graded as 1, 4/3, 2, 3, 4, 6, 8, 10, 12, 14, 16, and >16 ounces. Taking >16 ounces to be represented by 20 ounces, the average serving size for the 80 respondents was 7.9 ounces. This appears to be entirely consistent with the estimate from the Atkin survey (no formal statistical tests were performed)..

6.6 The effect of cooking fish

6.6.1 PBC loss due to various cooking methods

Zabik and Zabik (1999) reviewed the losses of PCBs during processing and cooking food, including the losses on cooking of fish. The studies and results they cite agree with the quantitative analysis by Wilson *et al.* (1998), who summarized the results of the various experimental studies on mass loss of PCBs (and DDT) during cooking of edible portions of fish. Wilson *et al.* (1998) indicate that their analysis showed that "baking, frying, broiling, boiling, smoking and microwaving all effectively reduce the concentrations of PCBs in fish tissue,"[24] and that they could not show any particular effect on the loss of mass during cooking due to the initial mass of PCBs in the raw fillet, fillet lipid content, or skin removal. They wrote distributions that were supposed to represent the loss of PCBs on cooking, but they arbitrarily extrapolated their distributions from the highest measured loss of PCBs to 100% loss.

[24] The analysis was actually for mass loss during cooking, not concentration change during cooking. These are substantially different, because of the substantial loss of mass of fish fillets (principally due to loss of water) on cooking.

Examination of the studies[25] upon which Wilson *et al.* (1998) based their analysis indicates that most studies examined very few fish. Many of the studies used just three fish for each combination of circumstances examined; and where estimates of standard deviation or individual fish data were available, they indicated that there was either large variation from fish to fish, or that the experimental uncertainties were substantial. The latter is indicated, at least for early studies, by the results Zabik *et al.* (1982) in which the apparent mass of PCBs increased during cooking, a physical impossibility — Wilson *et al.* (1998) and Zabik and Zabik (1999) omitted those results from consideration, as we do here. The total number of measurements (either individual pairs — with and without cooking — of fish or fish fillets, or, in a few cases, composite pairs) was 315, although individual measurements are provided in only one study. There are eight available published studies providing usable information, and when the 315 measurements are summarized and listed separately by published study, fish species, source of fish, and cooking method, there are 56 such combinations (see spreadsheet Cooking_effect.wb3, Appendix B.16).

To estimate mass loss from edible portions of fish during cooking, all eight available studies were analyzed. There were sufficient data to adequately examine three methods of cooking — baking, broiling, and frying (deep frying was combined with pan frying, since they were indistinguishable in the data, and because they were not distinguished in surveys associated with the Kalamazoo River). Most of the available measurements are of individual fish, but a few represent composites of fish. Many fish species were examined in the various studies — bluefish, carp, catfish, salmon, smallmouth bass, trout, walleye, and white croaker. No distinction can be drawn between them since no consistent difference was evident in the available data. However, there is sufficient evidence to show that the fraction of PCB mass retained is not constant for the same cooking method in the differing circumstances of the various studies (for example, different species of fish, source of fish, date, and detailed cooking method).

What is required is an estimate of the long-term average PCB mass reduction due to cooking of fish from the Kalamazoo, under the conditions of cooking practiced by eaters of these fish. To represent the uncertainty in this average for each cooking method, the results within each published study corresponding to a particular cooking method and fish species were combined to estimate the fraction of PCBs surviving as estimated by that study. The resulting values are shown in Table 6.14. To construct approximate uncertainty distributions, the values for each

[25] Armbruster *et al.* (1987), Armbruster *et al.* (1989), Cichy *et al.* (1979), Lee and Lee (1985), Puffer & Gossett (1983), Reinert *et al.* (1972), Skea *et al.* (1981), Smith *et al.* (1973), Trotter *et al.* (1989), Zabik *et al.* (1979), Zabik *et al.* (1982), Zabik *et al.* (1995a), Zabik *et al.* (1995b), and Zabik *et al.* (1996). The first four contain insufficient information for the required analysis, and were omitted, as was Zabic *et al.* (1982) as described in the text. Here we also omit Reinert *et al.* (1972), for lack of data on PCBs and Smith *et al.* (1973) because we could not obtain it, and use Skea *et al.* (1979) in place of Skea *et al.* (1981) because we could obtain the former but not the latter, and they appear to be discussion of the same data. We also add Schecter *et al.* 1998.

cooking method, with unity adjoined, were spaced uniformly in probability from 0 through 1, and piecewise linearly interpolated. The resulting uncertainty distributions are shown in Figure 6.7. The means of these distributions are 0.784, 0.720, and 0.679 respectively for baking, broiling, and frying.

Table 6.14 Fraction of PCBs remaining after cooking.

Cooking method	Study	Species examined	PCB fraction remaining [a]
Bake	Zabik *et al.* 1979	Trout	0.592
	Zabik *et al.* 1995b	Chinook Salmon	0.647
	Trotter *et al.* 1989	Bluefish	0.758
	Zabik *et al.* 1995a	Walleye	0.811
	Skea *et al.* 1979	Smallmouth Bass	0.836
	Zabik *et al.* 1996	Lake Trout	0.858
Broil	Schecter *et al.* 1998	Catfish	0.423
	Zabik *et al.* 1979	Trout	0.473
	Zabik *et al.* 1995b	Chinook Salmon	0.553
	Zabik *et al.* 1995a	Walleye	0.748
	Zabik *et al.* 1996	Lake Trout	0.836
	Skea *et al.* 1979	Brown Trout	1.000
Fry	Skea *et al.* 1979	Smallmouth Bass	0.260
	Puffer & Gossett, 1983	White Croaker	0.559
	Zabik *et al.* 1995b	Carp	0.673
	Zabik *et al.* 1995a	Walleye	0.854

[a] We give 3 significant digits here to reproduce exactly what was used in the analysis. The accuracy of the individual estimates from each of these studies is much lower.

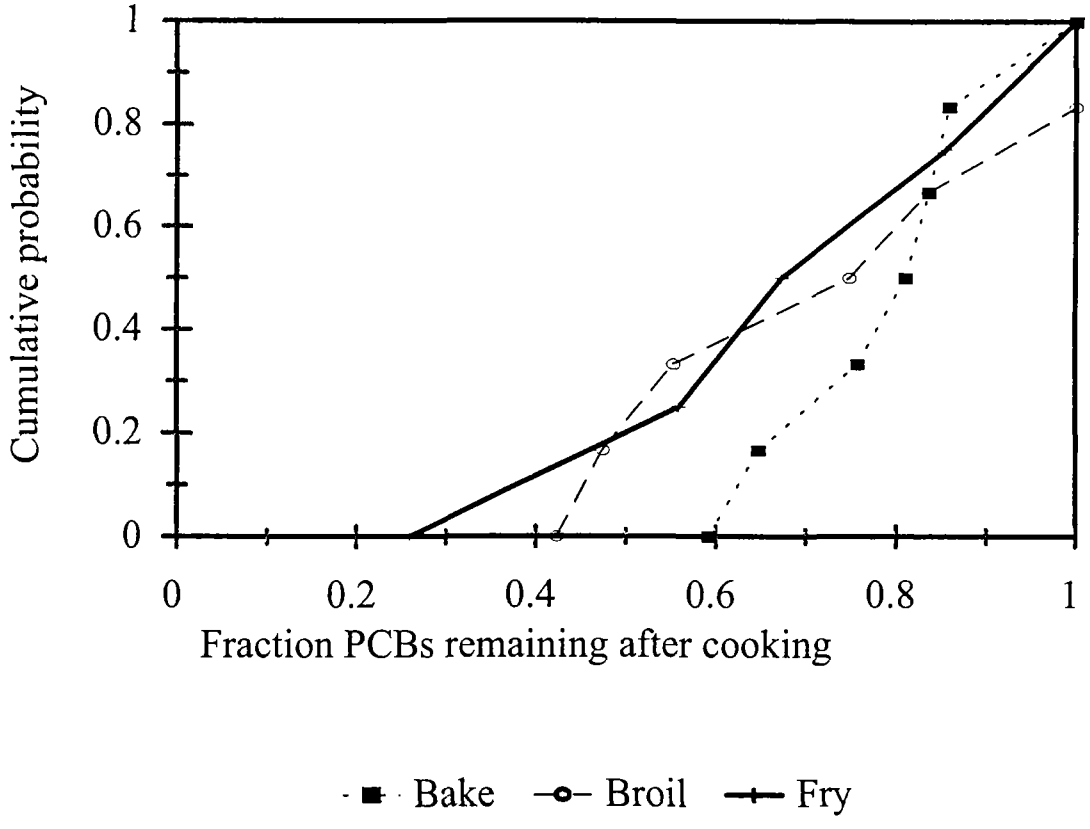

Figure 6.7 Cumulative uncertainty distributions for the fraction of PCBs remaining after cooking

6.6.2 Prevalence of cooking methods in this population

There is limited information available on the prevalence of various cooking methods in the population eating Kalamazoo fish. The first phase of the Kalamazoo River Angler Survey (MiCPHA, 2000a,b) did not contain any query on the type of cooking used. Phase II of that survey asked two questions: "When you eat sport-caught fish, how is it most often cooked", and got the responses (see spreadsheet Phase_2.wb3, Appendix B.18):

Broiled	6	Pan fried	44	Others	7
Baked	11	Deep fried	14	Unspecified	1

of 83 total responses. The follow-up question "if there is another usual method used to cook the sport-caught fish you eat, what is it?" obtained the responses

Broiled	12	Pan fried	11	Others	16
Baked	16	Deep fried	6	Unspecified	27

of the 88 total responses. The Atkin (1994) survey (see spreadsheet Atkin_survey.wb3, Appendix B.17) only asked one question about cooking methods for each type of fish eaten, "Do you fry the fish". The response was approximately:

Usually	70%
Sometimes	19%
Never	11%

when responses over all fish types examined are summed (the question was asked separately of bottom-fish like carps/suckers, smallmouth or largemouth bass, and other fish most frequently eaten; there were no significant differences in these relative response rates).

It appears that some form of frying, baking, or broiling is the most common cooking method, with most of the population usually preparing fish in such a way. There could, however, be a fraction of about 11% of the population that always uses stewing or some such cooking method that results in no loss of PCBs during cooking.

6.6.3 Variability and uncertainty distributions

The variability distribution for cooking methods in the fish-eating population is taken to be an 11% probability for long-term average cooking methods with no PCB loss, and an 89% probability for long-term average cooking methods with some PCB loss. The 11% figure is taken as a conservative estimate based on the "Never" response in the Atkin (1994) survey.

The various cooking methods examined appear to have similar cooking losses, but to incorporate the potential variability between individuals, we assume that the cooking loss for 75% of individuals is equivalent to that of frying, for 15% of individuals to the loss from baking, and for 10% of individuals to the loss from broiling. These are the relative fractions (rounded to the nearest 5%) in the Phase II Kalamazoo River Angler Survey (MiCPHA, 2000a,b), for these methods being the usual method of cooking (omitting other and unspecified categories), and are consistent with the results of the Atkin (1994) survey (see Section 6.6.2).

The uncertainty distribution for the long-term average loss from cooking by frying, baking, and broiling is evaluated using the uncertainty distributions given in Section 6.6.1. Using the

distributions described, the population average for the fraction of PCBs that survives any method of cooking is 0.732.

6.7 The population affected

There are no direct estimates available of the total population of anglers who eat the fish they catch on the Kalamazoo. Phase I of the Kalamazoo River Angler Survey (MiCPHA, 2000a,b) surveyed the population of anglers, but the available documentation omits aspects of its design and implementation (particularly on the levels of effort involved) that would assist in making such estimates. However, there is sufficient internal information available within that study to obtain rough estimates of the population, although the approach described here probably results in an estimate that is biased high. The calculations described in this section are performed in the spreadsheet Phase_1.wb3 (Appendix B.12).

The methods section of Phase I of the Kalamazoo River Angler Survey (MiCPHA, 2000a) describes the approach taken:

> MDCH made contractual agreements with the Allegan and Kalamazoo County publich health agencies to conduct the field portions of the survey. Each county fielded a two-person team consisting of a female and male interviewer. With personnel safety considerations in mind, surveying took place only during daylight hours. Additionally, members of the team always maintained visual contact with each other in the field.
>
> Pre-survey tours of the stream were made in both counties to identify locations where fishing was most frequently occurring and document the most popular hours/days of the week when fishing activity was highest. This information was taken into consideration in scheduling of survey team efforts during the most popular hours/days of the week. Field surveys were initiated in May 1994 and completed September 30, 1994.

Missing from available material are the scheduling of the survey teams, the pre-season survey information, and the sections of stream that were observed by the survey team during each session. 1090 anglers were observed, and 938 face-to-face interviews were conducted (MiCPHA, 2000a, although the database, MiCPHA, 2000b, contains 939 records) — the surveyors interviewed each individual at most once; but no records are available of the number of times each individual was met during the survey.

The database (MiCPHA, 2000b) shows that anglers were interviewed on 119 of the 151 days inclusive between May 3, 1994 and Sept. 30, 1994, the dates of the first and last interview recorded. Every day of the week was observed multiple times (Table 6.15); while there is a bias towards the weekends, weekdays were not omitted.

Table 6.15 Number of interviews by day of week, and number of weeks containing an interview on that day of the week.		
Day of week	Number of interviews	Number of weeks with an interview
Sunday	214	19
Monday	125	13
Tuesday	143	17
Wednesday	115	18
Thursday	67	16
Friday	99	17
Saturday	176	19

Entries for Allegan county indicated the location of the interview using the U.S. Public Land Survey system (township, range, and section). There were 65 unique locations identified, although up to 11 of these were unique by virtue of data entry errors. For the days on which interviews were conducted in Allegan county, those interviews were conducted in from 1 to 13 locations on a given day; and for every day of the week there was at least one occasion on which interviews were completed in at least 6 distinct locations. Thus each survey day was likely to have examined at least 10% of the river's length. However, it is not known how long individual anglers stay on the river on any given day, so the fraction of anglers observed on the survey day could be less than 10% of those present on the river at some time despite the planned bias towards popular locations.

Each angler was asked how often he or she had fished the relevant stretch of the Kalamazoo the last calendar year (Table 6.16), and in which seasons (spring, summer, fall, winter) he or she would fish the Kalamazoo (Table 6.17).

Table 6.16 Number of times the interviewed angler fished the Kalamazoo in the last calendar year.	
None	276
Once	43
2-5 times	172
6 or more times	444
Totals	659

Table 6.17 Seasons fished, by number of times fished the last calendar year, for Kalamazoo anglers.

Last calendar year	Fish in the season?	Spring	Summer	Fall	Winter
None	Yes	126	265	103	37
	No	149	10	168	233
	No response	1	1	5	6
Once	Yes	25	41	12	3
	No	18	2	30	39
	No response	0	0	1	1
2-5 times	Yes	100	156	67	17
	No	69	13	99	148
	No response	3	3	6	7
6 or more times	Yes	337	420	292	91
	No	105	24	150	349
	No response	2	0	2	4
Unknown	Yes	1	2	1	0
	No	1	0	1	2
	No response	2	2	2	2

With these data, we make an approximate estimate of the total population of anglers, and estimate an uncertainty range. We consider a simplified and idealized version of the survey, in which an angler visits the river on each of m days within the survey period of n (=151) days, each day choosing a random location. On each of p (=119) survey days within the survey period, the

surveyors randomly visit a fraction $1/k$ of the river (possibly in non-contiguous sections), with the assumption that the survey team will certainly observe the angler if he is within the fraction they survey that day (more generally, assume that the surveyors would certainly observe $1/k$ of all anglers on that day). The probability for an angler to be observed some time during the survey would then be

$$1-\left(1-\frac{1}{k}\frac{p}{n}\right)^{m} \tag{6.23}$$

This expression, averaged over the values of m available in the survey (1, 2–5, and 6 or more, interpreted as 6–12), was used to estimate a range of values by varying k from 2 (50% of anglers on the river surveyed on survey days) and 20 (5% of the anglers on the river surveyed on survey days), with a central estimate of 10. The central estimate, corresponding to 10% of the anglers on the river being observed on survey days, is derived from the observation above that each survey day was likely to have surveyed at least 10% of the river's length — and assuming that the targeting of popular spots compensates for the missing of anglers who arrived after or left before any survey sweep or sweeps along the river. This range is deliberately chosen somewhat wider than might be expected to account for uncertainties in the other estimates described below.

The values in Table 6.17 allow estimates for the probability of fishing in each season by annual number of visits. The survey questions on number of visits and seasons fished did not refer to the same time period, so these estimates will necessarily be approximate. For the purposes of this approximate calculation, we assume that those who did not visit in the last calendar year are similar to those who did. For those who visit the river only once, the probabilities for the four seasons must sum to unity; the observed probabilities do not because of the aforementioned disparity between the time periods — those who visited once during last calendar year might easily visit multiple times in other years. As an approximation, the probabilities for those who visited once last calendar year from the questionnaire are simply normalized to unity. Applying this approach, Table 6.17 leads to the first set of probabilities in Table 6.18.

The survey occupied one month in each of Spring and Fall, all three months of Summer, and omitted Winter entirely. For the Spring and Fall, if we assume that each month is equally likely to be fished, then for those fishing once the probability to be within the survey period in these seasons is 1/3 the probability for the whole season. For those fishing multiple times, the probability P_{ss} to have fished within the season and within the survey period is given approximately by[26]

$$P_{ss} = 1-\left(1-P_{s}\right)^{1/3} \tag{6.24}$$

[26] The approximation is in assuming that "multiple" times corresponds to a very large number of times. The approximation underestimates probabilities slightly, so results in slight overestimates for the population.

where P_s is the probability to have fished within the season. The resulting estimates are given as the set of probabilities in Table 6.18. For Spring and Fall, since the surveyed months were closest to the major fishing season, this procedure probably underestimates the probabilities, and hence biases our population estimates high.

Table 6.18 Probabilities to fish by season and number of times visiting.

Number of times	Spring	Summer	Fall	Winter
Probabilities to fish during the season				
None	0.46	0.96	0.38	0.14
Once	0.58	0.95	0.29	0.07
Once, normalized	0.31	0.50	0.15	0.04
2–5 times	0.59	0.92	0.40	0.10
6+ times	0.76	0.95	0.66	0.21
Probability to fish during the season and within the survey period				
Once	0.10	0.50	0.05	0
2–5 times	0.26	0.92	0.16	0
6+ times	0.38	0.95	0.30	0

The overall probability to fish within the survey period can now be estimated as the sum of the seasonal probabilities (for those fishing once), and approximately as

$$P_{survey} = 1 - \left(1 - P_{sp}\right)\left(1 - P_{su}\right)\left(1 - P_{fa}\right)\left(1 - P_{wi}\right) \tag{6.25}$$

for the other cases, where P_{survey} is the probability for fishing any time during the survey, and P_{sp}, P_{su}, P_{fa} P_{wi} are the probabilities for fishing during the survey in spring, summer, fall, and winter respectively.

The resultant calculation for anglers missed by the survey are shown in Table 6.19 for the best estimate (k=10). Each estimated number is the number in the survey divided by the product of the two probabilities shown.

Table 6.19 Adjustment of observed numbers to account for the incompleteness of the survey (best estimate).				
Number of times fishing during survey	Number in survey	Probability for observation by survey	Probability to fish during survey period	Estimated number
Once	43	0.08	0.66	831
2–5 times	172	0.25	0.95	733
6+ times	444	0.52	0.98	881
Total	659			2445
Ratio				3.71

Since the survey was entirely conducted during daylight hours. Any anglers who fish only at night, would have been missed by the survey. It seems unlikely that more than 10% of angler would fish only at night, so we increase the population estimate by this fraction. Since the observed number of anglers was 1090, the best estimate for the total number of anglers is $1090 \times 3.71 \times 1.1 = 4448$.

Not all anglers eat the fish, but others than the anglers also eat fish. The survey indicated that 0.469 of the anglers ate the fish they caught, but 3.29 additional persons ate fish per eating fisher, resulting in the best estimate of the fish-eating angler population of 6,870.[27] This is uncertain, however. We estimate the potential range range by examining the effect of varying k as described above. This gives a range from 2,401 to 12,793. We do not know the shape of this uncertainty distribution, so will encode it as triangular (with extremes given by the range just stated, and mode at the best estimate). In the sensitivity analysis (Section 6.11.1) we evaluate the effect of using a lognormal uncertainty distribution for the population size, with mode equal to 6,870 and standard deviation adjusted to allow 10% probability of exceeding a population size of 12,793 (the parameters of the resultant distribution are: median 7,909, logarithmic standard deviation 0.37527). Figure 6.8 shows the two assumptions for the uncertainty distribution. With the alternate, lognormal, uncertainty distribution there is approximately a 4.4% chance for the population to exceed 15,000.

[27] Between 8.5% and 17% of the anglers responding also indicated giving fish to friends, but there is no information on how many such friends. The question asked was whether "all fish" were given to friends, but the responses clearly indicate that the question was interpreted in alternate ways by some anglers. Any such fish transfers are here incorporated into the uncertainty of the fish-eating population.

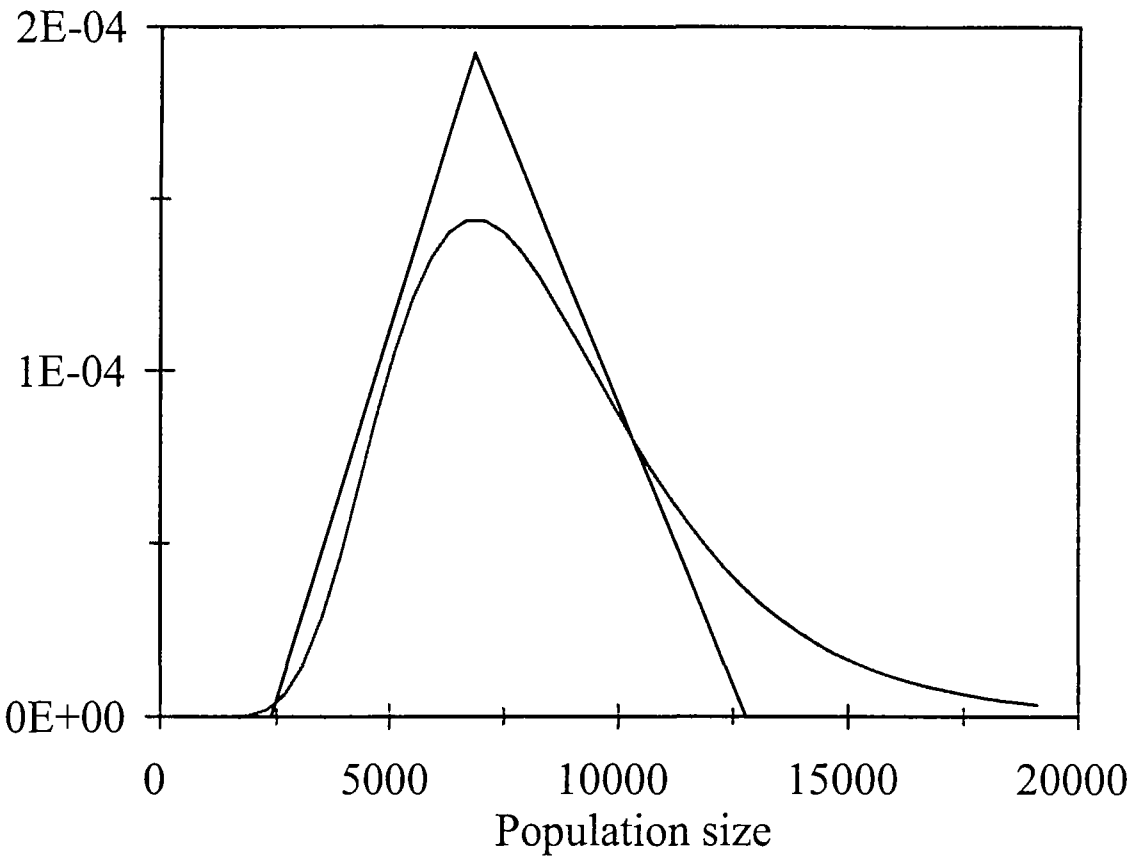

Figure 6.8 Alternate uncertainty distributions for population size. Triangular — base case; lognormal — alternate.

6.8 Evaluation of the total population effect

Evaluation of a total population effect requires accounting for the differences among the individual members of the population. Equation 6.3 gives the lifetime risk for individual members of the population who start eating fish in 1999, and during the Monte Carlo modeling we average over the population variability distribution to obtain the population average lifetime average dose rate, and hence a population average lifetime risk, R_{pop}, applicable to the population who start eating the fish in 1999 (there is an uncertainty distribution still associated with this population average). This risk estimate applies in aggregate to the population of anglers eating the fish they catch, and we apply it also to the additional population of persons eating fish

supplied by those anglers, although the upper ends of the distribution probably overestimate the risk for the total population who eat fish from the Kalamazoo River (including the anglers, their families, and others to whom they give fish). This total population, N, and its uncertainty, is evaluated in Section 6.7. The population turnover rate r is estimated using Equation 6.19, and it also has an uncertainty that is estimated during the Monte Carlo analysis.

The number of people entering the population of ever-eaters of Kalamazoo River fish is then rN per year, and the total population of ever-eaters alive at any one time is about $70rN$ (since the average lifetime is 70 years). For the rN people entering the ever-eater population in 1999, the expected number of cancers caused by PCBs from Kalamazoo Fish is rNR_{pop}, compared with a background rate of cancers of about $0.4rN$ (about 40% of people will get cancer from all causes during their lifetime; Ries *et al.*, 2001). For the rN persons entering the ever-eater population the following year, the expected number of cancers during their lifetimes is $rNR_{pop}\exp(-\beta)$ where β is the decay constant for PCB concentrations evaluated in Section 6.2.4. Similar decreases by a factor $\exp(-\beta)$ apply for all subsequent years.

We can approximate the pattern of expected cancer occurrences by assigning back to the year they start eating fish the expected probability of cancer for each member of the ever-fish-eating population, and adding up over all the people who start eating fish that year (although in truth any such cancers would occur many years after they started eating the fish). This approach ensures that we include every member of the ever-eating population in the calculation. Any such cancers would in reality be spread over the lifetimes of the total population who start to eat fish from the Kalamazoo at any time in the future.

Finally, the total population effect can be obtained by integrating the annual effect over all future years. The result is that the total number of cancers expected in the whole population over all future time is given by

$$n_{tot} = rNR_{pop}/\beta \tag{6.26}$$

This number is uncertain, of course; it has an associated uncertainty distribution due to the uncertainties in each of its terms, and we can estimate its uncertainty percentiles and its expected

value.[28] The total number of cancers expected in the whole population is less than one with high probability, and is generally not an integer. An alternative way of expressing this expected value is to write the probabilities for exactly no cancers, one cancer, two cancers, and so on. To a very good approximation the probability p_r for exactly r cancers, given an expectation of n_{tot}, in a large population is (Poisson approximation)

$$p_r = \frac{n_{tot}^r}{r!}\exp\left(-n_{tot}\right) \tag{6.27}$$

These probabilities have (highly correlated) uncertainty distributions, and we obtain their expected values by averaging over those uncertainty distributions. In particular, the probability for exactly zero cancers is obtained by averaging this expression with r=0 over the uncertainty distribution.

[28] The uncertainty distribution for the total number of cancers expected has a long right tail, but in reality must have finite expectation value because there is a finite total quantity of PCBs. In this analysis, the uncertainty distribution of β has been estimated as normal, truncated at a small, finite, lower bound. If instead that lower bound was zero, the expected value of the number of future cancers would be infinite, because the distribution of $1/\beta$ would then have infinite first moment (the infinity comes from the resultant logarithmic divergence of the expected value integral). A Monte Carlo simulation of such a situation will generally give an incorrect, finite, result — an estimate of infinity can only occur if the value zero for β is sampled, a highly unlikely event. Indeed, in a Monte Carlo simulation, the probability for sampling values of β low enough to make a difference in the estimate of the expected value is often very small, so a Monte Carlo simulation in such circumstances usually gives the same result as with a small, positive, lower bound on β, but may occasionally produce a wildly large value (or infinity) for the expectation value. In the present case, the physically derived lower bound on β (Section 6.2.4) is large enough that this problem does not arise.

The long right tail of the distribution has another effect — it is difficult to obtain an accurate estimate for the expected value because of the large standard deviation of the distribution, and hence large standard error of the mean in the Monte Carlo simulation, even with a large number of samples. The approximate lognormality of the final distribution for population effect is used to some advantage by computing a minimum variance unbiased estimator of the mean (Gilbert, 1987), assuming the distribution is lognormal.

6.9 *Unquantified uncertainties*

Some uncertainties have not been quantified for this risk assessment. The principal ones are as follows:

- As discussed in Section 4, the calculations are all conditional on PCBs actually harming human health at the dose rates evaluated here, which are very low compared with those in the laboratory rat experiments. For cancer, there is considerable doubt as to whether any carcinogenic effect would occur in humans or even in rodents at such very low dose rates. Even if it does, the application of a low-dose linear extrapolation from the ED_{10} must be considered a further unquantified uncertainty, since all the analyses performed in Section 4.2 are of high-dose experiments, and the interspecies extrapolation uncertainty of Crouch (1996) was derived for high doses. For non-cancer effects, there is no doubt that effects occur at some high enough dose rates. However, there is considerable doubt about whether the particular mild effects seen in monkeys are suitable for selecting a dose rate that can be extrapolated reliably to indicate a dose rate in humans that would result in adverse effects.

- Even with the assumption that PCBs actually cause cancer at low doses in humans, there are unquantified uncertainties in the method adopted for analysis of the bioassays. The variation between experimental results for Aroclors with the same designated name was assumed to be adequately represented by a variability between rat strains, even though it is known that there is variation in the congener profiles of Aroclors that are given the same name. A variation of toxicity due to differing congener profiles might also upset the assumption of constant relative potency for the different Aroclors within each rat strain. The experiment that dominates the relative potency estimates is Mayes *et al.* (1998) — only Schaeffer *et al.* (1984) provides any other contribution — so this uncertainty mostly hinges on the representativeness of the Aroclors used in Mayes *et al.* (1998). An extreme example of the problem of differing congener profiles is the assignment of Clophens A-30 and A-60 (Schaeffer *et al.*, 1984) in the analysis. As mentioned in the footnote to Table 4.1, these are assumed to be equivalent to Aroclors 1016 and 1260 respectively, based on their homolog fractions. An alternative assignment of Clophen A-30 as equivalent to Aroclor 1242 was tested, but found to be much more unlikely in the context of the relative potency model adopted. Thus there is considerable uncertainty in the relative potency estimates; but the sensitivity of the results to these assignments is not high — a factor two change in any of the relative potency estimates might change the overall result by +16%/–8% (see spreadsheet PCB_cancer_dose_response.wb3, Appendix B.5).

- The analysis of Section 4.2 was made on the assumption that the relative potencies of the Aroclors is the same for the same experimenter and rat strain. No attempt has been made to ensure that the results of the experiments are statistically significant — all that is required to obtain non-zero potencies in this approach is that overall, under this

hypothesis of equal relative potency, the maximum likelihood estimates for potency be non-zero. It is undoubtedly true that Aroclors 1254 and 1260 cause cancers in some experiments on laboratory animals; but it is by no means certain that Aroclors 1016 and 1242 do in any. We are extending the generic working hypothesis (our main unquantified uncertainty) that "PCB causes cancer" to all the Aroclors. Our failure to include the uncertainties in the ED_{10} estimates (particularly the possibility for an infinite ED_{10}) may therefore have resulted in a dramatic overestimate of the potency for these lower-chlorinated Aroclors; although the effect on this risk assessment would not be large, since the majority of the PCBs in fish are identified as higher chlorinated Aroclors (see, for example, Table 6.11).

- The analysis presumes that the fish concentrations of PCBs are adequately represented by the summary estimates of equivalent Aroclor concentrations. More precisely, it assumes that those summary estimates of Aroclor concentrations allow a match with the toxicity estimates for individual Aroclors. This uncertainty is currently unavoidable; even were complete congener concentration profiles in the fish available, the toxicity studies that indicate the most sensitive endpoints have been performed on Aroclors, not on individual PCB congeners.

- The characteristics of the fish-eating angler population is assumed to be in steady state, and accurately characterized by the Kalamazoo River Angler Survey (MiCPHA, 2000 a,b,c). It seems unlikely that there will be a large change in the population eating fish from the Kalamazoo. The survey asked some questions about expected increases and decreases in angling and fish-eating in various circumstances, but these have not been correlated with current fishing or eating habits. This survey continues to provide the best available estimates of the current and future fish-eating habits of the population.

- The fish concentrations are assumed to continue decreasing at a rate consistent with that observed over the last 10 years or so. This assumption is consistent with steady burial of the PCBs by sedimentation, but there are two possible circumstances that might disturb the trend. First, a major event that results in mixing of the deeper sediment to the surface, or substantial additional erosion of the former impoundment soils into the main river course, may increase the concentrations available to fish and hence result in a break in the steady time trend of fish concentrations. Major flooding or human intervention are the most likely potential causes. Second, if there is an up-river source of PCBs contributing to the contamination, the PCB concentration might not continue decreasing exponentially to zero, but instead decrease to a non-zero level. In this case our assessment excludes risks due to PCBs from the up-river source.

- There are several further biases built into the Monte Carlo analysis. The concentration data are slightly biased high, because of the approximation used for the distribution of mean concentration. The exposure periods (length of time of eating Kalamazoo fish) distribution was adjusted from the survey distribution to account for the effect of a one-

time survey on this distribution. No adjustment was made (on meal frequency, or length of time fishing, for example) for potential biases due to the survey having a higher probability to observe people who fish more frequently. The observed correlations between exposure period, meal frequency, and number of times fishing last season are small, but rather smaller than might be expected. The population size estimate is adjusted for this bias (see Section 6.7), and that adjustment implies a very low probability for observing people who fish the Kalamazoo only once per year (Table 6.19). It appears implausible that the population of those who fish the Kalamazoo only once per year, which makes up a substantial fraction of the estimated total population affected (approximately 1/3, see Table 6.19), could have fish consumption as large as implied by the risk assessment methodology.

- The assessment of the accuracy of the model (Section 6.11.2) indicates that on a measured (albeit self-selected) population of people fishing the Kalamazoo, using self-reported values for fish consumption and measured fish PCB concentrations, the risk assessment model may substantially overestimates (by a factor of about 7) the median dose to the population. The model is also unsuccessful in reproducing most of the variability in individual blood PCB concentrations, again based on the self-reported values for fish consumption and measured fish concentrations. These observations suggest that the risk assessment model may substantially overestimate doses to the individuals modeled, and hence to the population affected.

- There are always possible circumstances that are not considered in any risk assessment. We have not incorporated, for example, any scenario that involves deliberate eating of the soil from the former impoundments (or anywhere else).

6.10 Results of modeling

6.10.1 Variability across the population

Incorporating the analyses described in the preceding sections, and setting all the uncertainty distributions at their maximum likelihood estimates (MLE) or means, the variability distribution for lifetime average dose rate of total PCBs is given in Figure 6.9.[29] This curve describes the variation in lifetime average dose rates among a population of fish-eaters who start eating fish in 1999. The effect of later starting years is simply to multiply this distribution by the factor 0.953 per year, due to the exponential decay of the concentrations in fish. Approximately 93% of such a fish-eating population would have a lifetime average dose rate of 0.05 μg/kg-day or less, corresponding to a lifetime risk estimate of from zero to approximately 1×10^{-4} (assuming the

[29] This distribution was obtained from 1,000,000 Monte Carlo iterations. See spreadsheet Dose_life_results.wb3, Appendix B.21.

U.S. EPA range of cancer potency estimates — from zero to an upper bound of 2 kg-day/mg), 90% would have a lifetime average dose rate of 0.035 μg/kg-day or less, 95% would have a lifetime average dose rate of 0.071 μg/kg-day or less, 99% would have a lifetime average dose rate less than 0.24 μg/kg-day, and 99.9% would have a lifetime average dose rate less than 0.85 μg/kg-day.

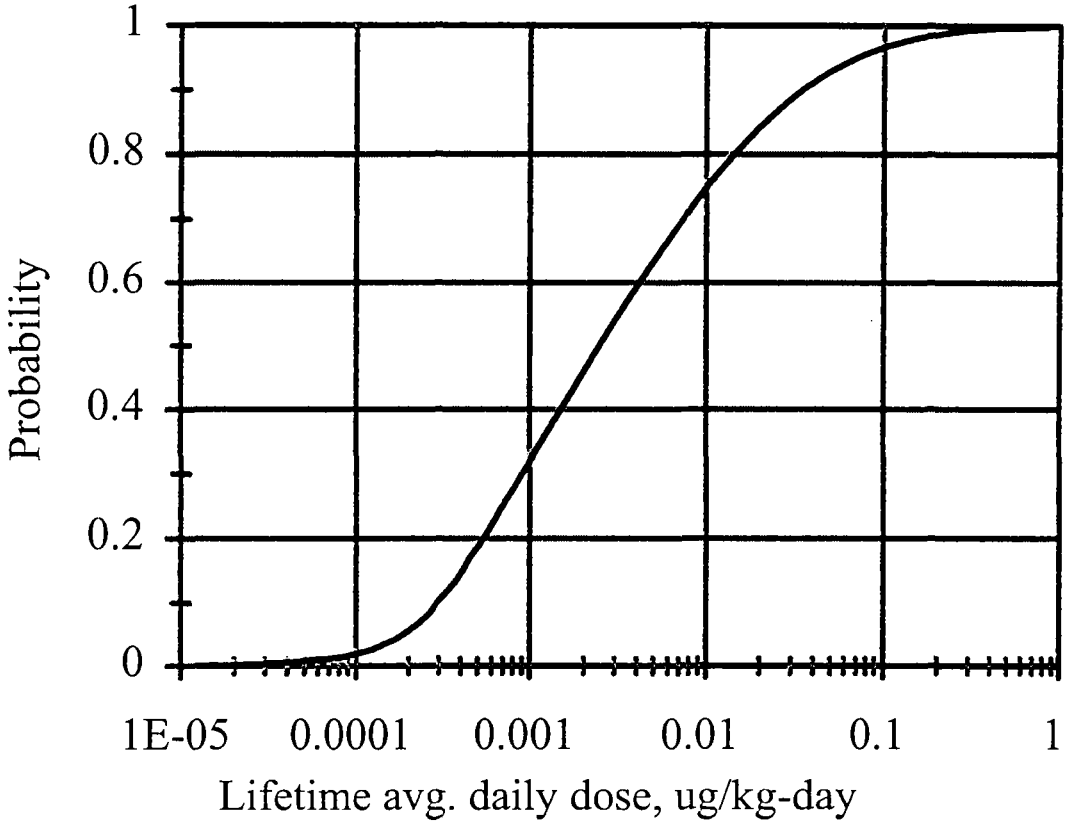

Figure 6.9 Population variability in lifetime average daily dose, μg/kg-day, with maximum likelihood estimates for uncertainty

The variability distribution for lifetime average daily dose is approximately lognormal[30] (the parameters of the best fitting lognormal are a median of 0.0029 μg/kg-day and a geometric

[30] This is so even though not one of the variability distributions or uncertainty distributions included in the calculations is lognormal.

standard deviation a factor of approximately 6.35). This distribution describes the differences among the population due to the differing habits of each individual — such as the differences in numbers of meals of fish eaten per year, the length of time for which they eat fish during a lifetime, and so forth. In principle, it would be possible to identify where any particular individual lies on this variability distribution by finding out for that individual how much fish he eats, for how long he eats fish from the Kalamazoo during his lifetime,[31] how large are his meals of fish, which fish species he eats, and where he catches them.

The distribution of dose rates averaged over the periods during their life that people actually eat fish from the Kalamazoo is shown in Figure 6.10 (again, this is for uncertainties set at central estimates — MLEs or means).[32] These dose rates are higher than the lifetime averages shown in Figure 6.9, because most people do not eat fish from the Kalamazoo for their whole lives. During the time they eat fish from the Kalamazoo, approximately 51% of the fish-eating population would have dose rates below the 0.05 μg/kg-day that was endorsed as safe for long-term exposure by the Michigan Environmental Science Board (Fischer *et al.*, 1998), while the 90th, 95th, 99th, and 99.9th percentiles are at 0.27, 0.45, 1.22, and 3.44 μg/kg-day respectively. However, these dose rates occur over periods ranging from 1 year to a lifetime, so that comparison with a single safe dose rate is problematic, as discussed in Section 4.3.1. In contrast, the hazard index results shown in Section 6.10.4 below account for the dosing period as well as the dose rate.

[31] This requirement to know for how long during a lifetime the individual eats fish from the Kalamazoo indicates that it would only be possible to identify the location on the variability distribution for any individual at the end of his lifetime.

[32] See spreadsheet Dose_while_results.wb3, Appendix B.22.

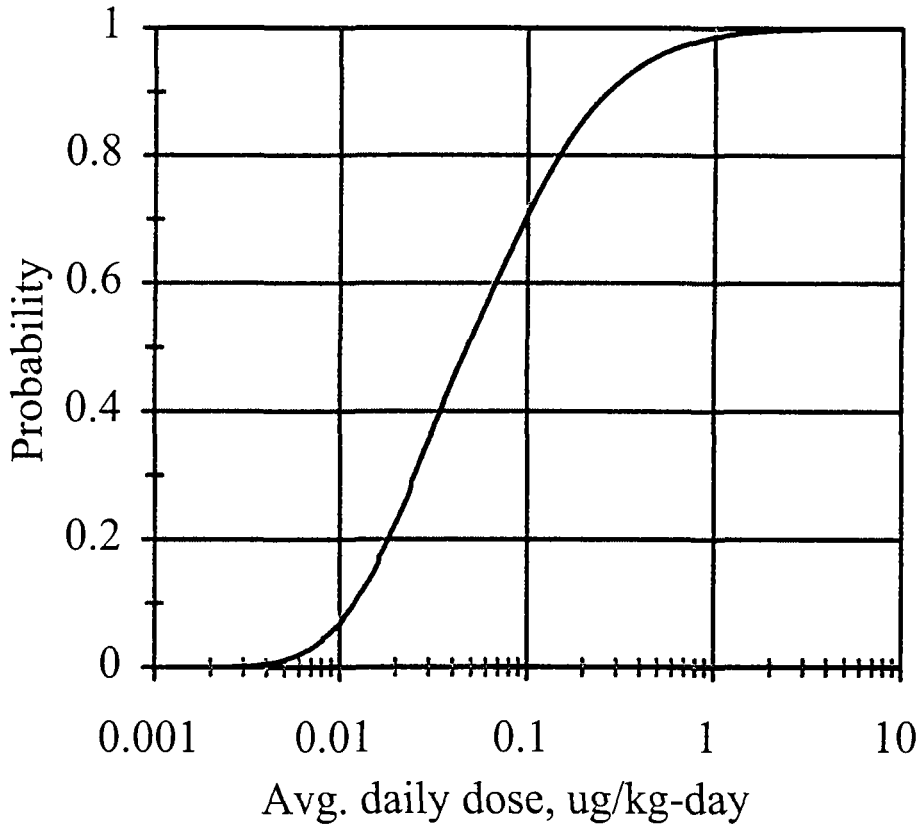

Figure 6.10 Population variability of the average dose rate (μg/kg-day) during the period of their lives that people actually eat fish from the Kalamazoo.

6.10.2 Uncertainties of the variability distribution

In addition to the variation in dose rates from individual to individual, there are uncertainties about the average dose rate for any individual. The uncertainties incorporated in the modeling have been described in individual sections above, and from them we have estimated the uncertainties associated with the variability distributions.

Incorporating all the identified uncertainties leads to uncertainty distributions for the variability distribution for doses described in Section 6.10.1. Figure 6.11 shows the distribution of uncertainties for the 50th, 75th, 90th, 95th and 99th percentiles of the variability distribution for

lifetime average dose rate.[33] For all the variability percentiles, the uncertainty distribution is fairly well represented by a lognormal with a geometric standard deviation of approximately 1.43. The horizontal line in Figure 6.11 shows the location of the MLE estimate for the variability distribution on these uncertainty distributions — the MLE estimate is at about the 25th to 40th percentile of the uncertainty distribution.[34]

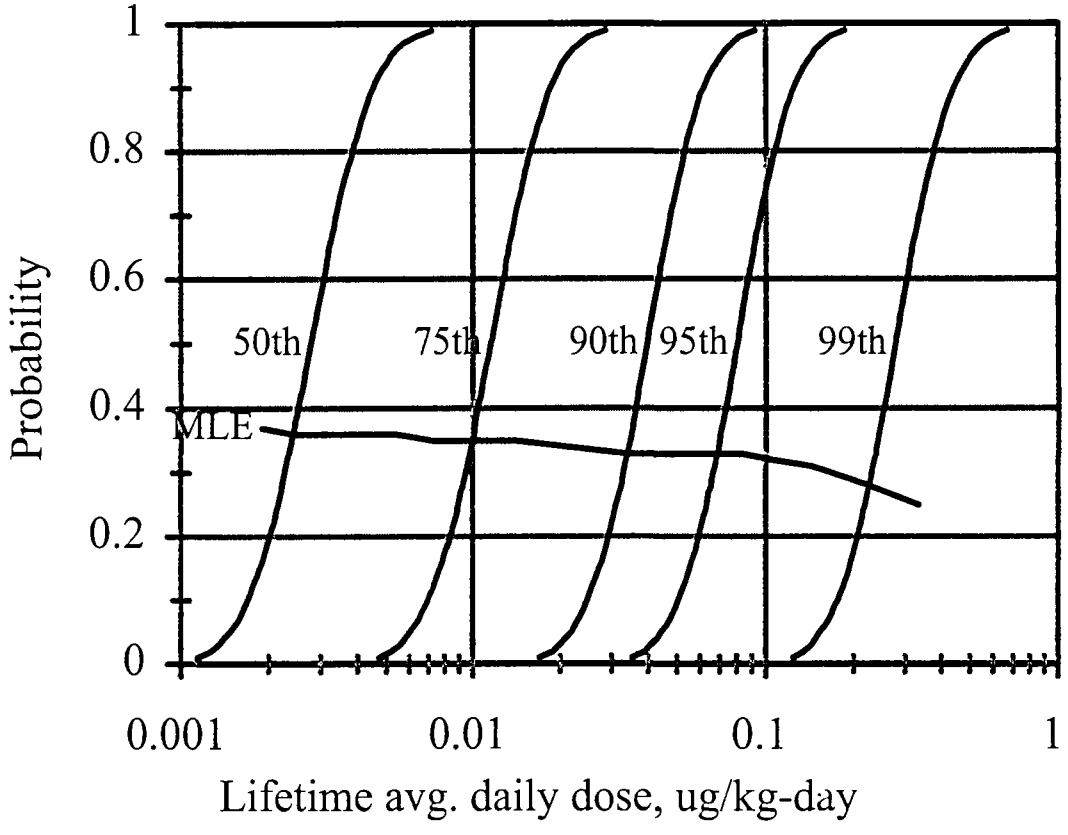

Figure 6.11 Uncertainty distributions for various percentiles of the variability distribution for lifetime average dose rate.

[33] See spreadsheet Dose_life_results.wb3, Appendix B.21.

[34] The jaggedness of the horizontal line is largely an artefact. The Monte Carlo simulation was performed using 50,000 iterations for the variability distributions, repeated 5,000 times with different samples to obtain the uncertainty distributions. The position of the MLE was evaluated only to the nearest 1 percent in positioning this line.

For the 90th percentile of the variability distribution (MLE value 0.035 μg/kg-day), the upper 90th percentile of the uncertainty distribution is at 0.062 μg/kg-day, corresponding to an estimate of lifetime risk estimate of between zero and 1.2×10^{-4}, assuming the fixed U.S. EPA upper-bound estimate of potency of 2 kg-day/mg. For the 95th percentile on the variability distribution (MLE estimate 0.071 μg/kg-day), the upper 90th percentile of the uncertainty distribution is 0.12 μg/kg-day, corresponding to an estimate of lifetime risk of from zero to 2.5×10^{-4} with the U.S. EPA upper-bound potency estimate. For the 99th percentile on the variability distribution (MLE estimate 0.24 μg/kg-day), the upper 90th percentile of the uncertainty distribution is 0.44 μg/kg-day, corresponding to an estimate of lifetime risk of from zero to 8.7×10^{-4} with the U.S. EPA upper-bound potency estimate.

Figures 6.12 and 6.13 both show, on slightly different scales, the full variability distribution for lifetime average dose together with its uncertainty. Figure 6.12 shows the MLE variability distribution (solid line on the left), together with (moving to the right) the 50th, 75th and 95th percentiles of uncertainty distributions about the variability distribution. Figure 6.13 shows the same, but with an inverse normal scale on the left — the straightness of the curves illustrates how close to lognormal is the variability distribution.

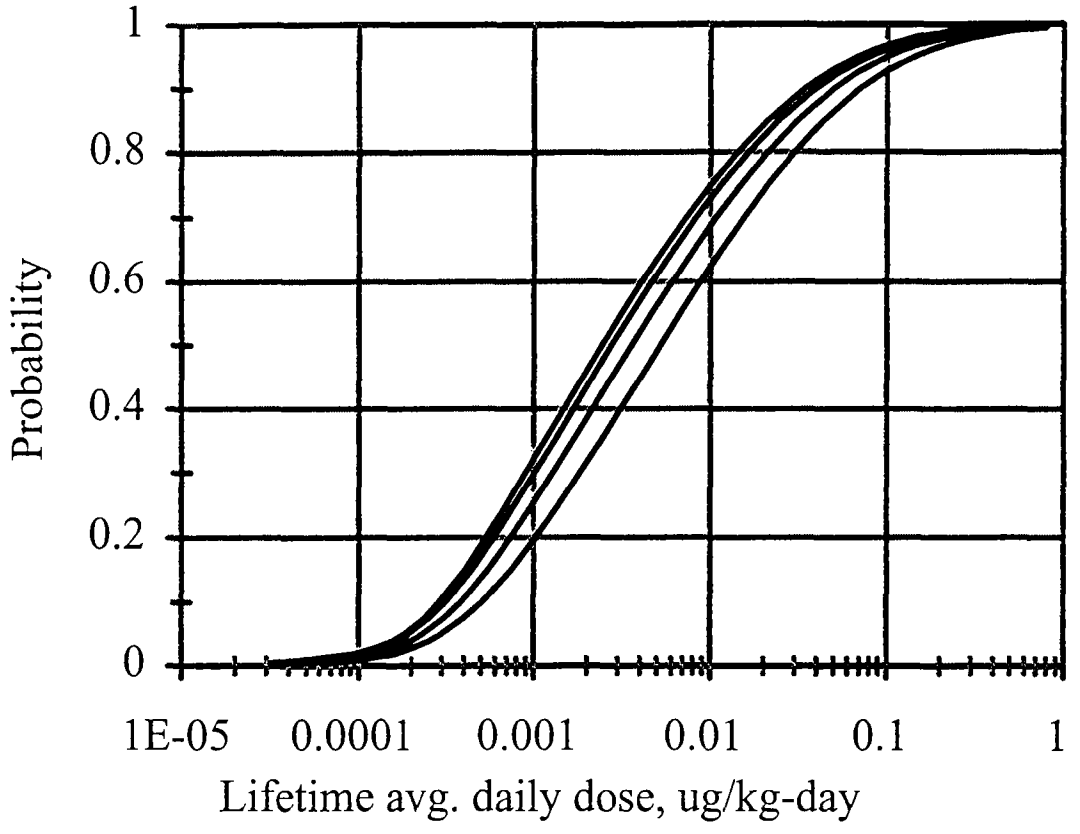

Figure 6.12 MLE variability distribution for lifetime average dose rate (to the left), and 50th, 75th and 95th percentiles (moving to the right) in uncertainty for this variability distribution.

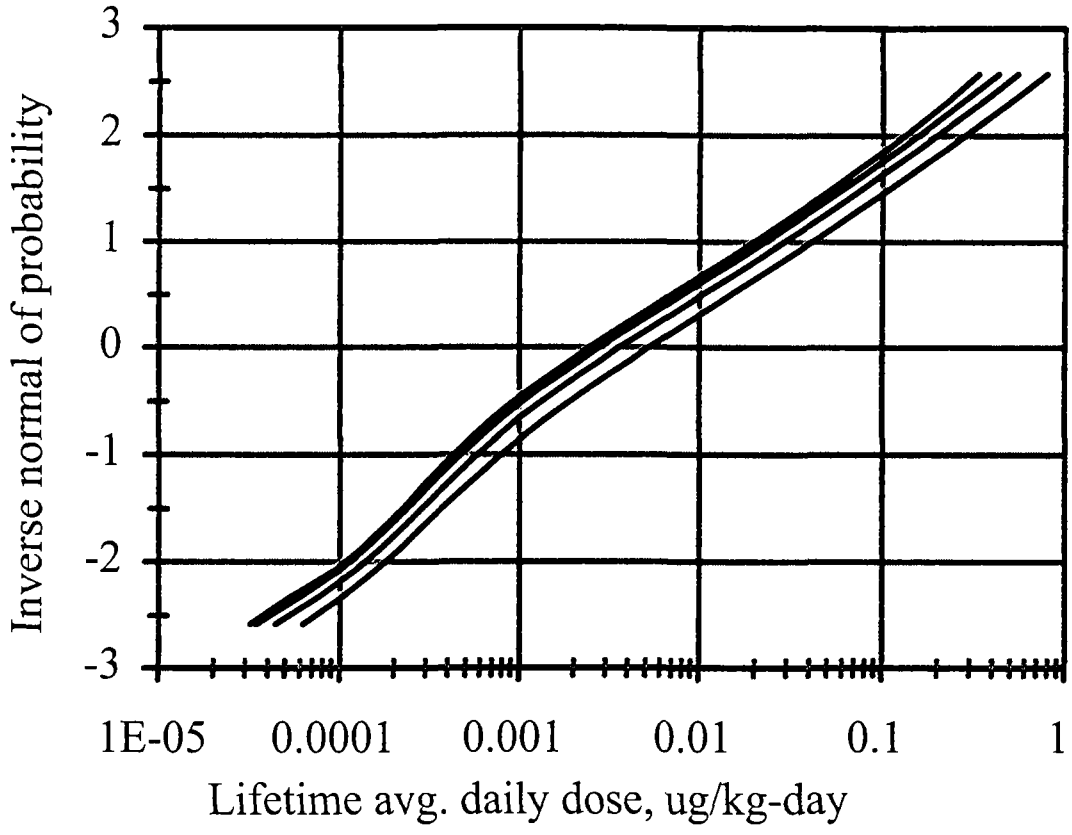

Figure 6.13 MLE variability distribution for lifetime average dose rate (to the left), and 50th, 75th and 95th percentiles (moving to the right) in uncertainty for this variability distribution (alternate scale).

The same type of analysis may be performed for the dose during exposure.[35] Figure 6.14 shows the distribution of uncertainties for the 50th, 75th, 90th, 95th and 99th percentiles of the variability distribution for average dose rate during exposure. For all the variability percentiles, the uncertainty distribution is fairly well represented by a lognormal with a geometric standard deviation of approximately 1.34. As before, the horizontal line in Figure 6.14 shows the location of the MLE estimate for the variability distribution on these uncertainty distributions — the MLE estimate is again at about the 25th to 40th percentile of the uncertainty distribution.

[35] See spreadsheet Dose_while_results.wb3, Appendix B.22.

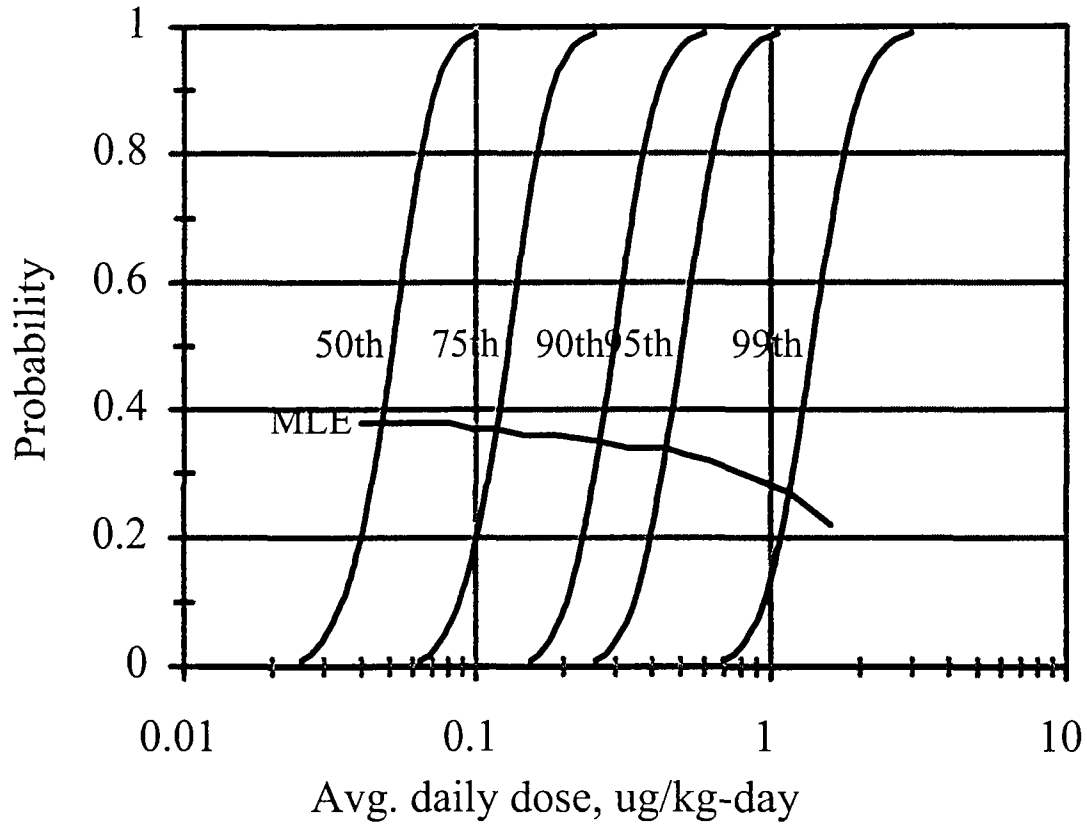

Figure 6.14 Uncertainty distributions for various variability percentiles of the dose during exposure.

For the 90th percentile of the variability distribution for dose during exposure (MLE value 0.27 µg/kg-day), the upper 90th percentile of the uncertainty distribution is at 0.42 µg/kg-day. For the 95th percentile on the variability distribution (MLE estimate 0.45 µg/kg-day), the upper 90th percentile of the uncertainty distribution is 0.72 µg/kg-day. For the 99th percentile on the variability distribution (MLE estimate 1.22 µg/kg-day), the upper 90th percentile of the uncertainty distribution is 1.98 µg/kg-day. Once again, however, comparison of these values with a single safe dose rate is problematic, as discussed in Section 4.3.1, since they occur over widely varying periods of exposure.

6.10.3 Combined variability and uncertainty — the random individual

For a randomly chosen individual, about whose habits we know nothing except that he eats fish from the Kalamazoo, there is no distinction between variability and uncertainty — the selection at random makes the variability equivalent to uncertainty. For such a randomly chosen individual, the uncertainty distribution for lifetime average dose rate may be obtained from the modeling by treating variability and uncertainty equivalently. This is the usual situation for uncertainty modeling, and corresponds to the practice in most risk assessments (including the HHRA) of simply choosing values from the various variability and uncertainty distributions without regard to whether they reflect variability or uncertainty. Performing this evaluation leads to the combined distribution for lifetime average dose shown in Figure 6.15.[36] This is almost indistinguishable from the variability distribution shown in Figure 6.9, because the uncertainty is so much less than the variability. The 90th percentile is at 0.041 µg/kg-day, the 92nd percentile is at 0.05 µg/kg-day, the 95th percentile is at 0.084 µg/kg-day, the 99th at 0.30 µg/kg-day, and the 99.9th at 1.09 µg/kg-day. This combined distribution is well approximated by a lognormal (with parameters of: median 0.0032 µg/kg-day, geometric standard deviation a factor of 6.77).

[36] See spreadsheet Dose_life_results.wb3, Appendix B.21.

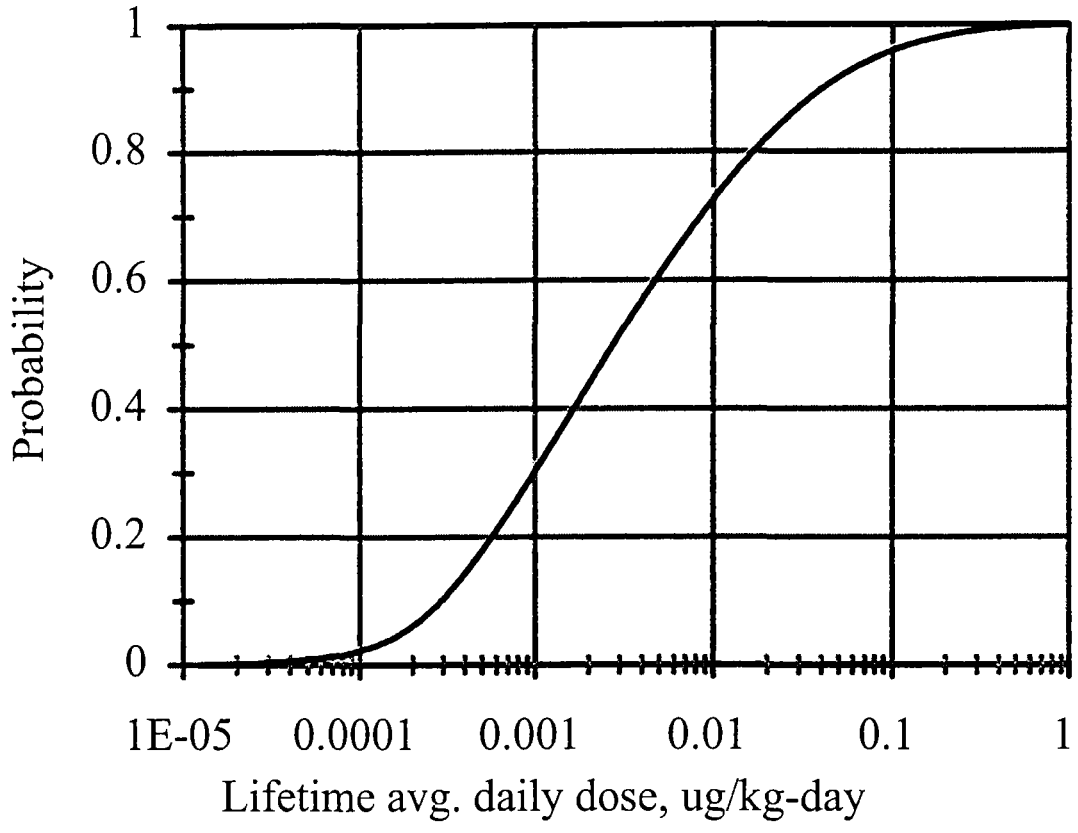

Figure 6.15 Combined variability and uncertainty for lifetime average dose rate — the uncertainty distribution for a randomly picked individual.

There is a similarly small effect of uncertainties on the estimates of dose rate during exposure (so that the graph, Figure 6.16, is almost indistinguishable from Figure 6.10) — again the variability is much larger than the uncertainty.[37] Approximately 49% of people randomly selected from the fish-eating population would have dose rates below the 0.05 μg/kg-day that was endorsed as safe for long-term exposure by the Michigan Environmental Science Board (Fischer *et al.*, 1998), while the 90th, 95th, 99th, and 99.9th percentiles are at 0.31, 0.53, 1.5, and 4.5 μg/kg-day respectively. Once again, these average doses occur over periods ranging from one year to a lifetime, so that comparison with any single acceptable level is problematic (see Section 4.3.1).

[37] See spreadsheet Dose_while_results.wb3, Appendix B.22.

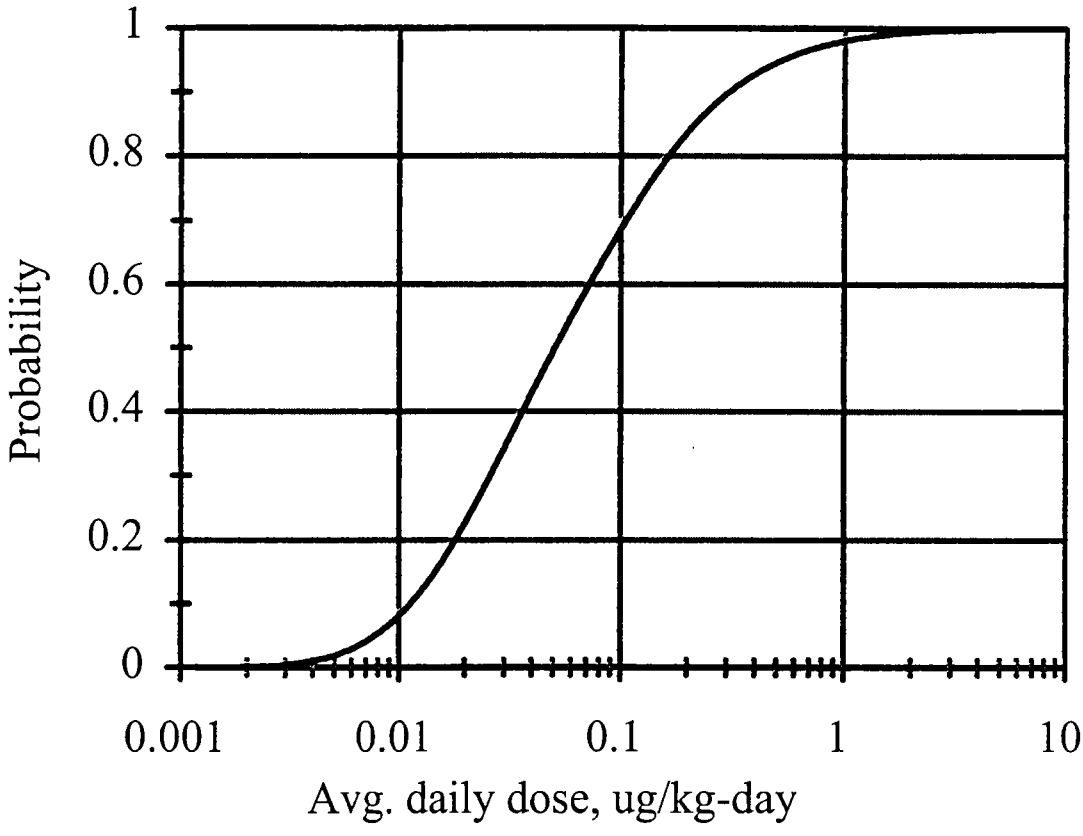

Figure 6.16 Combined variability and uncertainty for average dose rate during exposure — the uncertainty distribution for a randomly picked individual.

6.10.4 Results incorporating toxicity uncertainties

Sections 6.10.1, 6.10.2, and 6.10.3 describe results that take account of the uncertainty and variability in dose estimates, but do not account for the variability and uncertainty of toxicity values. As described in Sections 4.2 and 4.3, we have evaluated the latter variabilities and uncertainties, conditional on PCBs having such effects on humans at all, and incorporated them into the calculations. Figure 6.17 shows the uncertainty distributions for the 50th, 75th, 90th, 95th,

and 99th percentiles of the variability distributions for lifetime risk.[38] There are very large uncertainties involved, even conditional on the assumption that PCBs cause human cancers at all, so that the uncertainty distributions extend over a very wide range of risks. For example, at the 90th variability percentile, the 10% to 90% range of uncertainty is from 3.3×10^{-7} to 1.0×10^{-4}, almost 300-fold.

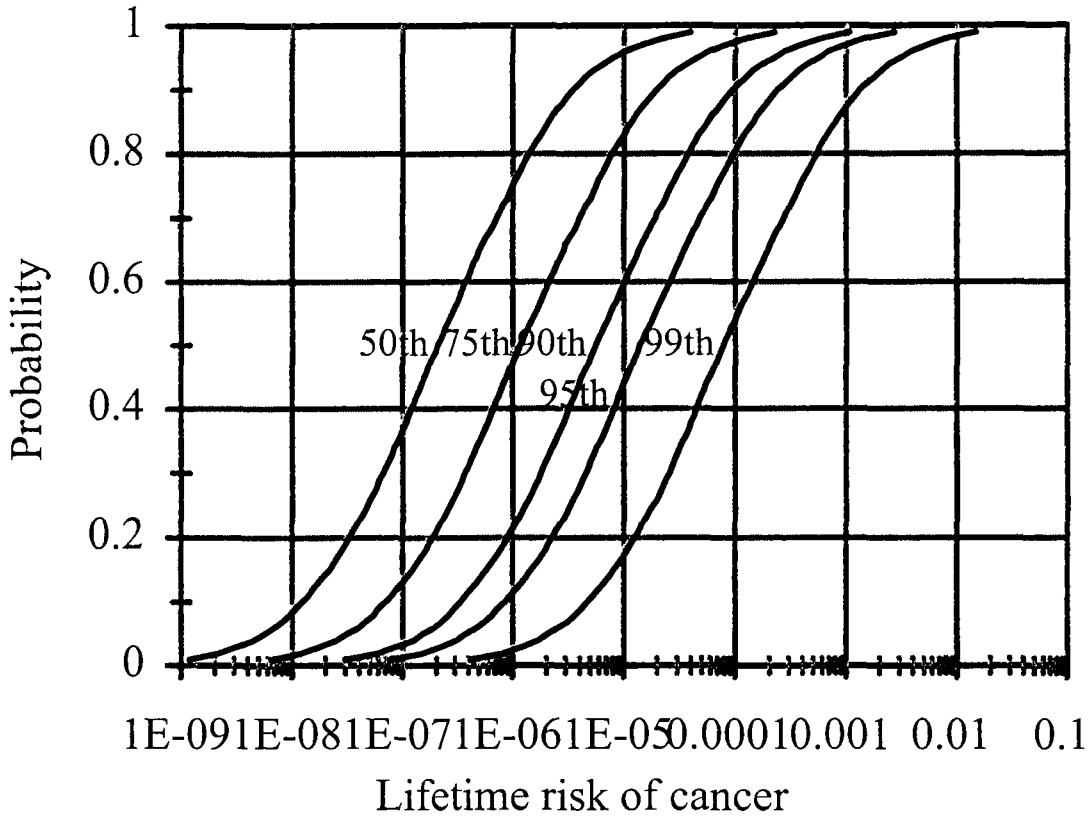

Figure 6.17 Uncertainty distributions for percentiles of the variability distribution for lifetime risk.

For a randomly chosen member of the fish-eating angler population, we can combine the variability and uncertainty distributions, since the random selection converts the variability into an uncertainty. This corresponds more closely to the usual calculations performed in EPA-style risk assessments, where parameter values supposed to be representative of the RME individual

[38] See spreadsheet Risk_results.wb3, Appendix B.23.

are selected on the basis of both variability and uncertainty distributions. The uncertainty distribution for lifetime risk for such a random member of the fish-eating population is shown in Figure 6.18. The potential risk ranges from completely negligible values below 1×10^{-6} with 67% probability, with a 13% chance that it exceeds 1×10^{-5}, a 3.6% chance that it exceeds 1×10^{-4}, and about 0.7% chance that it exceeds 1×10^{-3}. For this distribution, the 50th percentile (median) is at 2.2×10^{-7}, the 90th percentile is at 1.7×10^{-5}, the 95th percentile is at 5.8×10^{-5}, and the 99th percentile is at 5.7×10^{-4}.

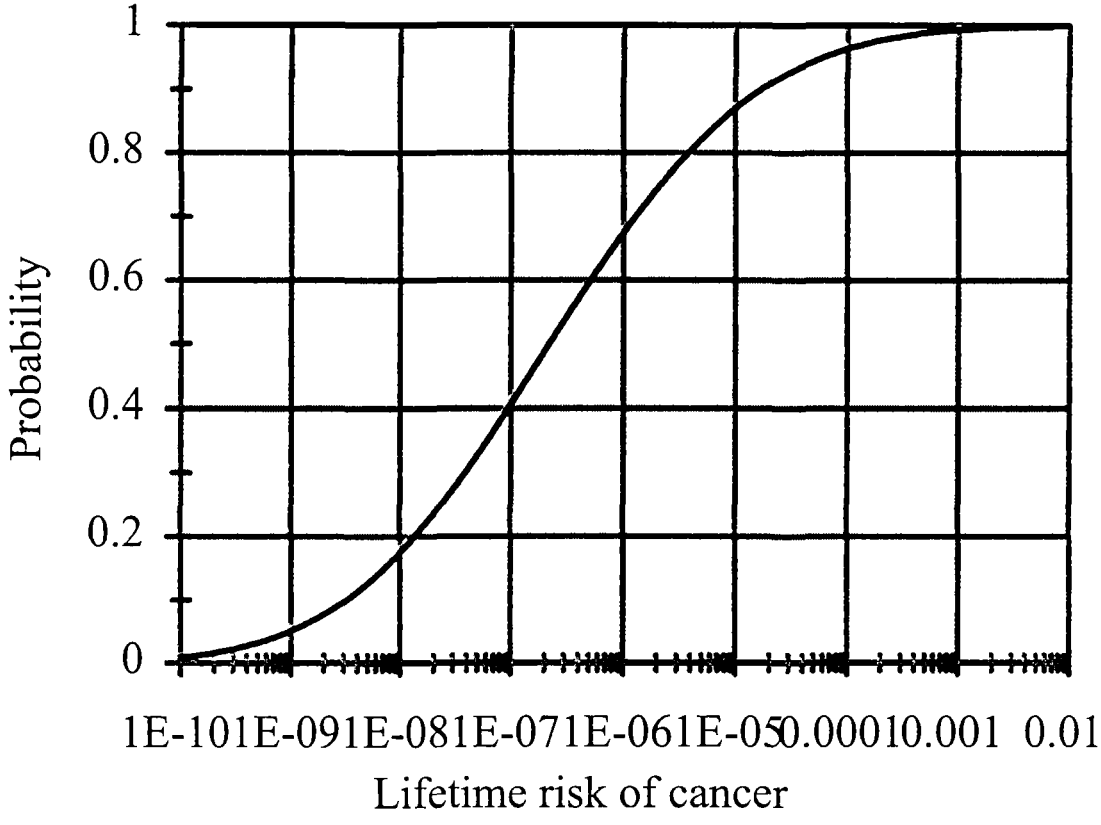

Figure 6.18 Uncertainty distribution for lifetime risk for a randomly chosen fish-eating angler.

To illustrate the combinations of circumstances that lead to lifetime risk estimates of 1.0×10^{-5}, at the 87th percentile of the combined uncertainty and variability distribution, Table 6.20 shows a selection of ten equally likely possibilities (these were taken from the Monte Carlo simulation; they are from the ten simulations giving risk estimates closest to 1.0×10^{-5}). The average PCB concentration listed in Table 6.20 is an average of the PCB concentrations in different fish

species, weighted by the fraction of meals of those species, and averaged over the period of exposure. Table 6.21 shows the corresponding 10 combinations of average PCB concentrations in 1999 for the individual fish species, and the fractions of meals of each of those species. The concentrations shown in Table 6.20 may be obtained from Table 6.21 by weighting the fish concentrations by the meal fractions, and then accounting for the decline with time of the PCB concentration — see Equations 6.1 and 6.7. Table 6.20 also includes a column for the carcinogenic potency of Aroclor 1254 — these examples incorporate the uncertainties in the potency of PCBs, so the potency is different for each example. However, the effective potency applicable in each line differs (and differs from the value given for Aroclor 1254), because the analysis takes account of the different Aroclor mixtures in each fish species (Table 6.9), the mix of fish species eaten (Section 6.5.2), and the relative potencies of the different Aroclors (Section 4.2.4), as in Equation 6.3 (see spreadsheet Examples.wb3, Appendix B.20, for a detailed calculation check of all the examples given in this section, and others).

Table 6.20 Examples of combinations of circumstances that result in a risk estimate of 1.0×10^{-5}. (see spreadsheet Examples.wb3, Appendix B.20)

Initial age (years)	Duration eating fish (years)	Effective additional duration (years)	Fish meals per year	Average weight of a meal (kg) [a]	PCB survival during cooking	PCB conc. decrease per year [b]	Average PCB conc. (mg/kg)	Potency of Aroclor 1254 (kg-day/mg)
30.8	5.1	0.6	2.95	0.34	1.000	0.0586	7.03	0.61
27.5	8.2	1.1	2.43	0.23	0.572	0.0500	1.53	5.83
57.7	6.4	-2.5	2.96	0.34	0.778	0.0403	2.21	3.85
7.8	22.8	5.1	25.99	0.34	1.000	0.0546	0.66	0.16
24.6	23.8	1.8	52.44	0.23	0.846	0.0212	0.86	0.13
47.4	4.9	-0.5	2.64	0.26	0.928	0.0427	2.23	4.24
22.8	2.0	0.4	176.02	0.11	0.818	0.0455	1.41	0.47
16.4	1.0	0.2	6.00	0.23	0.843	0.0649	0.94	18.90
21.9	3.1	0.6	23.89	0.29	0.706	0.0282	0.95	1.53
13.9	1.0	0.3	1.99	0.34	0.908	0.0511	4.03	7.84

[a] The fish meal weights in this column correspond to those in Table 6.13. For example, 0.34 kg = 12 oz, 0.23 kg = 8 oz.

[b] The decrease per year in the natural logarithm of the concentration.

Tables 6.20 and 6.21 show combinations of circumstances corresponding to a lifetime risk of 1.0×10^{-5}. It is apparent that a wide range of combinations of circumstances can lead to the same estimates of risk — it is impossible to focus on just one or two circumstances as being the major contributors. To extend the examples and emphasize this point, Table 6.22 illustrates combinations of circumstances that correspond to a risk ten times higher, at 1.0×10^{-4}.

Table 6.21 Fraction of meals of each species of fish, together with average concentration in those fish in 1999, for the ten examples in Table 6.20. Each entry shows the fraction of meals above the concentration in mg/kg. (see spreadsheet Examples.wb3, Appendix B.20)

Walleye	Sucker	Carp	Bass	Pike	Panfish	Catfish	Turtle
0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
0.812	1.938	8.127	0.841	2.666	0.373	1.041	0.309
0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
0.914	1.595	2.837	1.459	2.970	0.337	1.866	1.101
0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
0.742	1.858	2.033	2.335	2.010	0.317	2.510	0.804
0.008	0.000	0.000	0.060	0.008	0.200	0.724	0.000
0.680	1.538	2.495	0.985	7.561	0.353	1.325	0.822
0.393	0.000	0.000	0.071	0.036	0.000	0.143	0.357
0.951	0.772	1.808	1.847	2.270	0.479	1.244	0.912
0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
1.160	0.726	4.085	1.480	3.098	0.464	2.473	0.863
0.023	0.000	0.000	0.159	0.068	0.068	0.682	0.000
0.615	0.553	3.308	1.488	3.537	0.390	1.401	1.152
0.093	0.006	0.031	0.593	0.016	0.105	0.151	0.005
0.913	1.724	5.374	0.654	2.785	0.459	1.507	0.468
0.056	0.185	0.000	0.148	0.000	0.185	0.370	0.056
0.658	1.366	2.948	1.298	2.802	0.390	1.012	1.066
0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
0.676	0.783	4.132	1.016	2.891	0.400	1.013	0.596

Table 6.22 Examples of combinations of circumstances that result in a risk estim 1.0×10^{-4}.(see spreadsheet Examples.wb3, Appendix B.20)

Initial age (years)	Duration eating fish (years)	Effective additional duration (years)	Fish meals per year	Average weight of a meal (kg) [a]	PCB survival during cooking	PCB conc. decrease (fraction per year)	Average PCB conc. (mg/kg)	Potency of Aroclor 1254 (kg-day/mg)
14.9	8.6	1.9	49.36	0.11	0.886	0.0498	0.68	7.24
14.4	1.1	0.3	15.90	0.34	0.404	0.0534	1.23	70.82
17.9	4.6	1.0	16.72	0.34	0.932	0.0651	1.99	4.39
26.6	5.4	0.8	26.13	0.34	0.757	0.0378	0.27	24.37
51.5	4.9	-0.9	93.82	0.23	0.772	0.0594	1.03	3.90
55.7	2.3	-0.6	5.23	0.11	0.829	0.0270	0.92	327.51
31.6	9.2	0.8	79.44	0.23	0.610	0.0293	1.30	1.78
9.7	6.3	1.7	139.21	0.34	0.705	0.0556	0.67	1.39
9.3	30.6	5.5	314.63	0.34	0.534	0.0499	0.24	0.56
7.7	10.0	2.6	32.40	0.34	0.635	0.0539	0.83	3.18

[a] The fish meal weights in this column correspond to those in Table 6.13. For example, 0.34 kg = 12 oz, 0.23 kg = 8 oz.

For non-cancer effects, as for lifetime risk, there is a wide variability and uncertainty. Figure 6.19 shows the uncertainty distributions for fixed percentiles of the variability distribution.[39] It indicates that there is a small probability (less than 4%) that less than 1% of the population would receive a dose exceeding an individual no-adverse-effect-level by a factor of 100, and a probability of about 33% for less than 1% of the population to exceed an individual no-adverse-effect-level by a factor of 10. For 90% of the population, however, there is better than 70% probability that their doses will not exceed individual no-adverse-effect-levels.

[39] See spreadsheet HI_results.wb3, Appendix B.24.

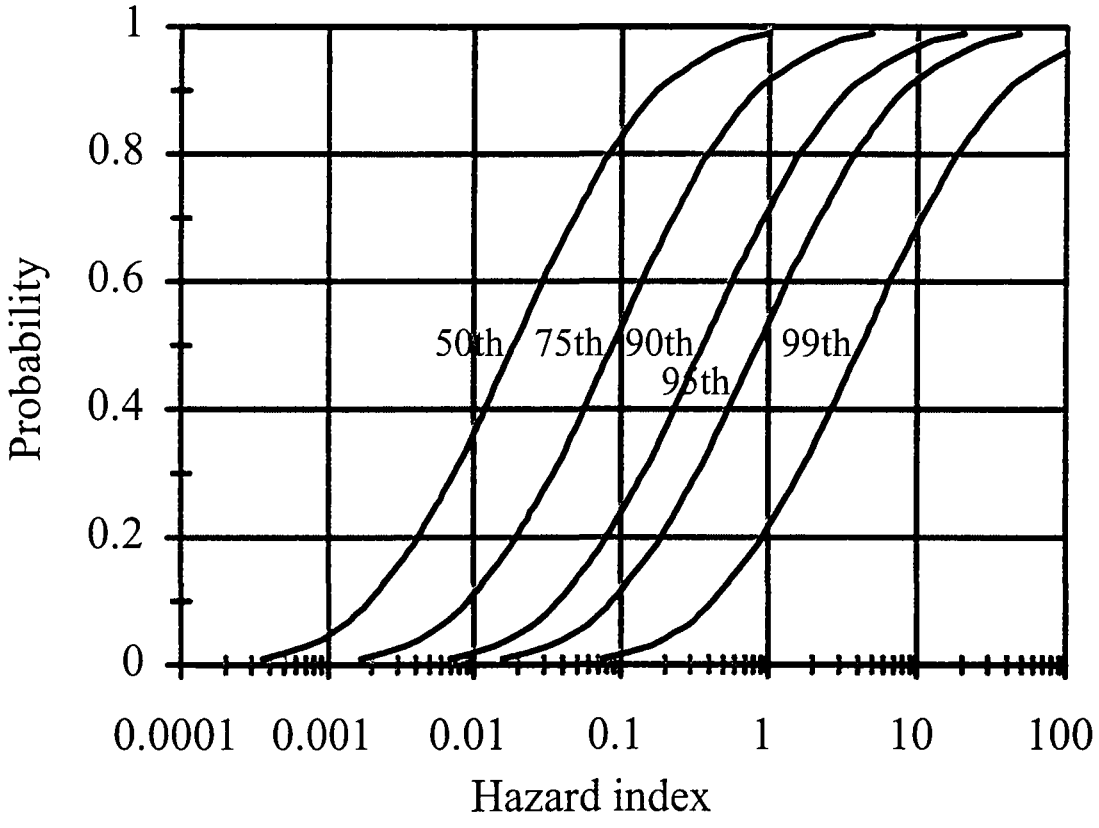

Figure 6.19 Uncertainty distributions for percentiles of the hazard index variability distribution.

Just as for the uncertainty distribution for lifetime risk, the uncertainty distribution for the hazard index for a randomly chosen individual has a very wide range. Figure 6.20 shows that there is approximately an 41% chance that such a randomly chosen individual will have a hazard index less than 0.01, 71% chance for less than 0.1, and 91% chance for less than 1. The probability (for a randomly chosen individual known to eat fish) to exceed unity, and so be at risk for some non-cancer effect, is thus approximately 8.9%, and the probability to exceed a hazard index of 10 is about 1.7%. This uncertainty distribution for hazard index for a randomly chosen individual has median (50th percentile) at 0.020, 90th percentile at 0.81, 95th percentile at 2.4, and 99th percentile at 17.

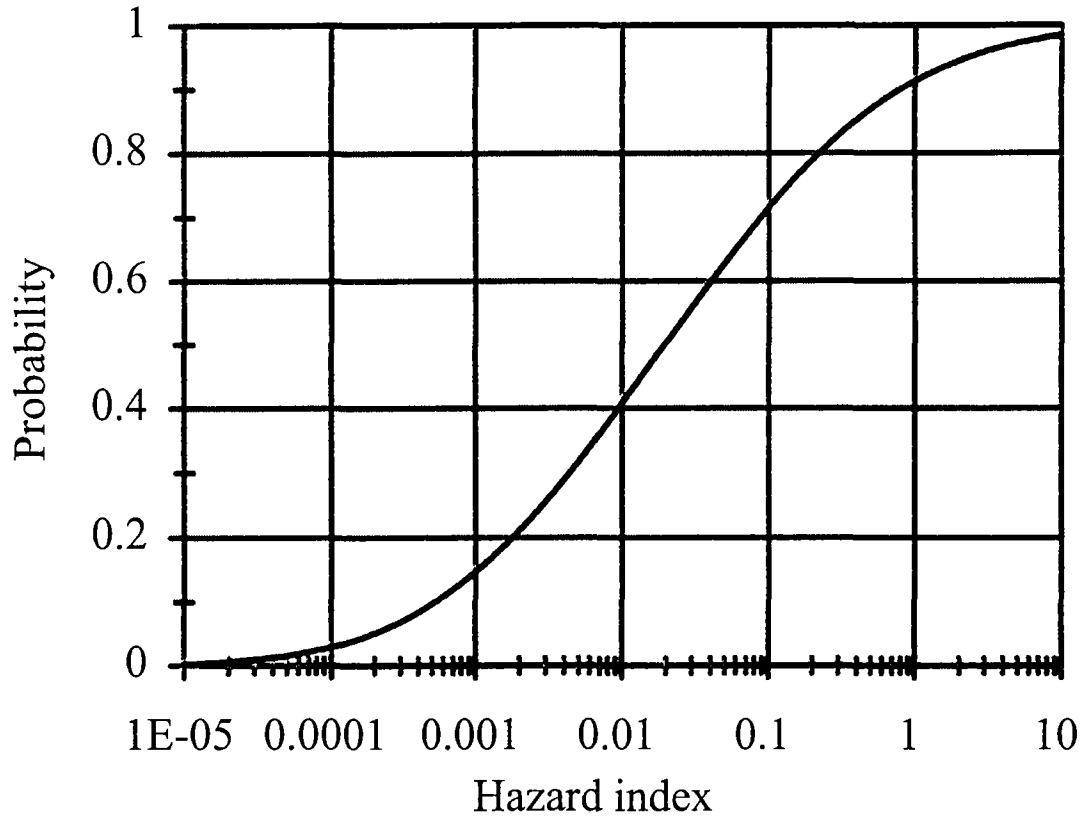

Figure 6.20 Uncertainty distribution for hazard index for a randomly chosen individual.

Again as for lifetime risk, the combinations of circumstances that lead to particular estimates of hazard index are legion. To illustrate, Table 6.23 shows ten such combinations of circumstances that lead to a hazard index estimate of 1.0, at the 92nd percentile of the combined uncertainty and variability distribution (these were taken from the Monte Carlo simulation; they are from the ten simulations giving hazard index estimates closest to 1.0).

Table 6.23 Examples of combinations of circumstances that result in a hazard index of 1.0. (see spreadsheet Examples.wb3, Appendix B.20)								
Initial age (years)	Duration eating fish (years)	Effective additional duration (years)	Fish meals per year	Average weight of a meal (kg) [a]	PCB survival during cooking	PCB conc. decrease (fraction per year)	Average PCB conc. (mg/kg)	NOAEL for PCBs (µg/kg-d)
14.4	6.2	1.5	18.0	0.11	0.533	0.0440	1.75	0.074
32.0	3.5	0.4	9.9	0.34	1.000	0.0639	1.10	0.144
9.4	4.5	1.2	5.3	0.23	0.635	0.0590	3.90	0.116
2.9	35.5	7.3	49.8	0.34	0.565	0.0505	2.17	0.814
34.4	15.4	0.2	163.7	0.11	0.840	0.0663	1.32	0.803
46.0	8.2	-0.9	29.8	0.11	0.881	0.0581	1.51	0.176
31.9	2.9	0.3	19.7	0.11	0.732	0.0412	2.77	0.177
13.2	2.1	0.5	8.6	0.34	1.000	0.0433	1.49	0.171
18.6	4.4	0.9	36.1	0.23	0.605	0.0332	1.50	0.290
17.1	1.0	0.2	43.0	0.34	0.715	0.0402	1.07	0.437

[a] The fish meal weights in this column correspond to those in Table 6.13. For example, 0.34 kg = 12 oz, 0.23 kg = 8 oz.

6.10.5 Population effect

Evaluation of a total population effect requires accounting for the differences among the individual members of the population. The Monte Carlo approach we have taken allows us to do this by averaging over the variability distribution to obtain the population average for lifetime average dose, allowing estimation of the total population effect, as explained in Section 6.8. At the MLE for uncertainties, the mean value for lifetime average dose in the population of those eating fish is 0.021 µg/kg-day, corresponding to a lifetime risk estimate ranging from zero to approximately 4.1×10^{-5} (using the U.S. EPA upper-bound potency estimate of 2 kg-day/mg) for those entering the population in 1999 (the index year for these calculations).[40] This strictly applies to the population of anglers on the Kalamazoo who eat the fish they catch, although the upper bound estimate probably overestimates the values for the total population described in

[40] See spreadsheet Dose_life_results.wb3, Appendix B.21.

Section 6.7 who eat fish from the Kalamazoo River (including the anglers, their families, and others to whom they give fish). That total population number is on the order of 6,870 persons actively eating fish at any one time (see Section 6.7). The turnover rate is about 15% of that population per year (see Section 6.4), or about 1,004 persons/year; so that number enter the total population of ever-eaters of Kalamazoo fish. It follows that the population of ever-eaters that is alive at any one time is about 70,300 persons (1,004 persons/yr × 70 yr lifetime).

Using the approach of Section 6.8, the median estimate for the upper-bound long-term-average population effect of PCB contamination in fish from the Kalamazoo thus calculated is about 0.038 cancers per year among ever-eaters (based on those starting to eat the fish in 1999), using the U.S. EPA upper-bound potency estimate of 2 kg-day/mg. Again, based on the U.S. EPA upper-bound potency estimate, the expected number of cancers per year (the average over the uncertainty distribution) is 0.041, and the upper 90th percentile is 0.064. These may be compared with a background cancer rate from all causes of about 400 per year in the population of ever-eaters of Kalamazoo fish. The value would decrease by about 5% per year as the PCB concentrations decrease, leading to a best estimate of the upper-bound effect over all time (that is, adding up all the cancers that might occur due to the PCBs among all the people who ever eat fish from the Kalamazoo) of about 0.79 total cancers (median estimate — the mean and 90th percentile estimates are 1.0 and 1.7 respectively). Any such cancers would be spread over the lifetimes of the total population who start to eat fish from the Kalamazoo at any time in the future.

The values so far discussed used the 2 kg-day/mg. upper-bound carcinogenic potency estimated by the U.S. EPA. Taking full account of the variability and uncertainty in the potency estimate increases the estimated values slightly at the upper probability tail, and decreases them at the lower end of the distribution (Figure 6.21).[41] For cancers/year due to PCBs, the full uncertainty distribution has a 50th percentile, mean, and 90th percentile of 0.0053, 0.070, and 0.094 respectively. The uncertainty distribution for total number of cancers has median, mean, and 90th percentile 0.11, 1.5, and 2.2 respectively,[42] and the probability for one or more cancers is 26%. Overall, then, it is highly likely that no cancer will ever occur in the whole population due to consumption of PCBs in fish from the Kalamazoo.

---

[41] See spreadsheet Risk_results.wb3, Appendix B.23.

[42] The mean estimate was computed using a minimum variance unbiased estimator (MVUE), assuming the underlying distribution was lognormal — a very close approximation. Its corresponding MVUE standard error estimate (this is numerical uncertainty, due to the finite number of Monte Carlo iterations) is 0.1. The straight mean estimate is 1.8 with numerical standard error estimate 0.4. These large standard error estimates occur because of the extreme right-tailed nature of the distribution — reducing them appreciably by increasing the number of Monte Carlo iterations is impractical.

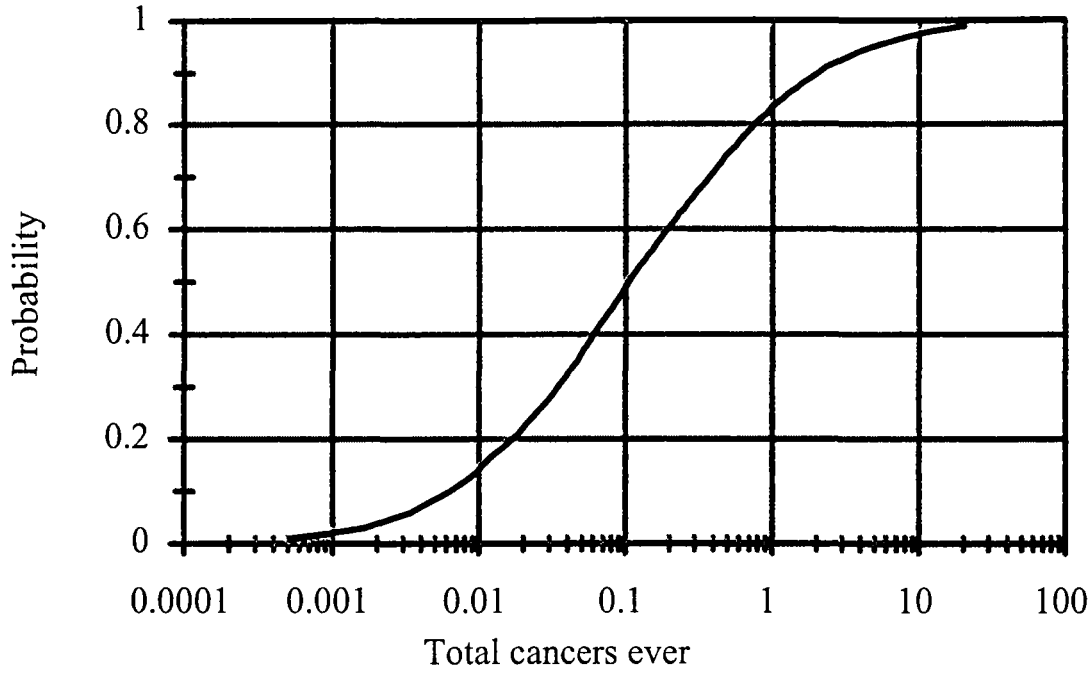

Figure 6.21 Uncertainty distribution for the total number of cancers ever occurring due to PCBs in fish

6.11 Sensitivity and accuracy of the model

6.11.1Sensitivity

By performing all the calculations of the Monte Carlo model while omitting various of the uncertainties discussed above, we can evaluate the sensitivity of the model to the uncertainties involved.[43] The dominating uncertainties in the results for lifetime risk and hazard index are the uncertainties in the toxicity values. The total uncertainty in lifetime dose estimates and dose estimates during exposure is about a factor 1.43 and 1.34 respectively, compared with uncertainties in toxicity values of at least a factor of 5 (geometric standard deviations in both

[43] See spreadsheets Dose_life_results.wb3 and Dose_while_results.wb3, Appendices B.21 and B.22 for the calculations of this section.

cases).[44] The majority of the following analysis was therefore performed on the dose estimates alone, to evaluate the much smaller effects of individual terms. To speed the analyses, they were performed using the Monte Carlo method with 10,000 iterations for determination of the variability distributions, repeated 1,000 times for the uncertainty evaluations (25 times fewer than the 50,000/5,000 used for reporting the results). This results in numerically less stable estimates for individual percentage points of the distributions, and experience with multiple computations indicates that the averages used in this sensitivity analysis fluctuate by as much as ±5%, but this is adequate to show the general picture.

An approximate measure of the relative importance of each uncertainty may be obtained by evaluating its contribution to the variance of the overall uncertainty estimate in doses. Since the uncertainty distributions for doses are approximately lognormal, we examined the contribution to the estimated variance of the logarithms. To obtain the estimates given in Table 6.24, we averaged the variance contributions of the logarithms of the uncertainty distributions over approximately equal numbers of the 31 saved percentiles (0.5%, 1.0%, 2.0%, 3.0%, 4.0%, 5.0%, 7.5%, 10.0%, 15.0%, 20.0%, 25.0%, 30.0%, 35.0%, 40.0%, 45.0%, 50.0%, 55.0%, 60.0%, 65.0%, 70.0%, 75.0%, 80.0%, 85.0%, 90.0%, 92.5%, 95.0%, 96.0%, 97.0%, 98.0%, 99.0%, 99.5%) of the variability distribution.

The contributions change over the variability distribution, as is to be expected. Apparently negative contributions are the result of numerical variability in the simulations (we are subtracting two values with relatively large uncertainties, and all runs used different random number sequences). Similarly, the "Total" contribution varies from 100%, probably reflecting some combination of the numerical variability, the complexity of the model (with all the interactions), and the use of the variance of the logarithms — the uncertainty variances do not quite just add up, but interact with each other and with the variability.

At the upper end of the distribution, the five uncertainty sources included in the table contribute approximately equally to the lifetime average dose rate. For the estimates of dose during exposure, the uncertainty in PCB losses in cooking increases slightly in relative importance, while the uncertainty in the lifetime period eating fish becomes negligible (not surprisingly, since the only effect this variable has on the dose during exposure is through the small correlation with meals per year). The "numerical uncertainties" correspond to the uncertainties inherent in the Monte Carlo process due to using a finite number of samples. Even with the smaller numbers used for this sensitivity analysis (10,000 samples for the variability distribution, repeated 1,000 times for the uncertainty estimates), the numerical uncertainties are small.

[44] See spreadsheets Dose_life_results.wb3 and Dose_while_results, Appendices B.21 and B.22 respectively

Table 6.24 Contributions to uncertainty in lifetime dose and dose during exposure.				
Uncertainty in:	Contribution to uncertainty variance			
	$0.5 - 10^{th}$ percentile	$15^{th} - 50^{th}$ percentile	$55^{th} - 90^{th}$ percentile	$92.5^{th} - 99.5^{th}$ percentile
Lifetime dose				
Lifetime period eating fish	24%	34%	38%	14%
Average meals per year	-3%	6%	6%	16%
Fraction of PCBs lost in cooking	52%	38%	30%	26%
Time trend of PCB concentrations in fish	-3%	3%	6%	20%
Fish concentrations	8%	18%	18%	14%
Numerical	2%	0%	1%	2%
Total	80%	100%	99%	93%
Dose during exposure				
Lifetime period eating fish	-0%	-2%	-1%	-3%
Average meals per year	-0%	7%	11%	21%
Fraction of PCBs lost in cooking	81%	63%	52%	34%
Time trend of PCB concentrations in fish	2%	6%	10%	15%
Fish concentrations	4%	23%	26%	24%
Numerical	1%	0%	0%	2%
Total	87%	97%	99%	93%

A further sensitivity analysis was performed, to evaluate the effect of inclusion or exclusion of the effective additional exposure period discussed in Section 6.3.2 on the estimates of lifetime risk (this additional exposure period is not used in the calculation of hazard index). Setting the extra exposure period to zero for any exposure duration resulted in a distribution of lifetime risk estimates that differed very little from the distribution obtained when the extra exposure period was included. The largest effect was at the very low end of the distribution (risk estimates below 1×10^{-8}), where the risk estimates for percentiles below the 50^{th} were increased by the omission

of the extra period by 5% at the 10th percentile to 30% at the 0.1th percentile (see graph time-effect in spreadsheet Risk_results.wb3, Appendix 20). At percentiles above the 50th, the differences between the distributions was generally within the numerical uncertainty of the Monte Carlo procedure — using 1,000,000 iterations, this uncertainty was generally about ±1% (standard deviation) at percentiles less extreme than the 98th.

The final sensitivity analysis performed was to the shape of the population uncertainty distribution, as discussed in Section 6.7. The alternate specification for the population uncertainty distribution, which has a long right tail and also implies a median population estimate about 7,909, approximately 15% higher than in the baseline case, gives slightly higher estimates (around 10%) for total number of cancers (Table 6.25, see spreadsheet Risk_results.wb3, Appendix B.23).

Table 6.25 Estimates for total population effect for baseline and alternate population uncertainty distributions.

Total cancers	Baseline	Alternate
Median estimate	0.11	0.12
90th percentile	2.2	2.4
Expected value	1.5	1.7

The probability for zero cancers in the whole population is also practically unchanged at 72% (versus 74% for the baseline case). While the population uncertainty is large, we conclude that plausible assumed shapes for the uncertainty distribution have little effect on the risk estimates made here.

6.11.2 Accuracy

The model used in this analysis of fish eating from the Kalamazoo River is that generally used for such risk assessments. The analysis of Section 6.10.2 shows that the precision of the model is high using the data available for the population of eaters of fish from the Kalamazoo — the uncertainties involved are generally a factor of about 1.4, which is very small compared with the variability between people. The accuracy of the model, however, is also open to question — that is, given knowledge about an individual, how accurate is the model at estimating that individual's PCB intake? Phase II of the Kalamazoo River Angler Survey (MiCPHA, 2000a,b) allows just such an assessment of the accuracy of the model, since the respondents in Phase II provided information just like that used in the model (for how long and how often they had been eating fish from the Kalamazoo, the types of fish, and their age), and an objective measurement — blood level of PCBs — was obtained that can be related to PCB intake.

In order to assess whether the consumption of fish from the Kalamazoo River has caused elevated levels of PCBs in humans, the State of Michigan performed a two-phase survey of people who fished along the affected part of the rivers (MiCPHA, 2000a). The first phase was designed to determine the number of the anglers using the river, what fish they caught, and how much they consumed or gave to others to consume. The second phase was a follow-up on a subset of participants in the first phase and included the measurement of PCBs in the blood serum of anglers who consumed fish and those who did not. We here examine data from the second phase of the survey (MiCPHA, 2000a,c) to determine whether there is a difference between the blood serum PCB levels of those who ate fish from the river, and those who did not.[45] Additionally the blood levels of anglers who ate fish are compared with predicted levels based on the anglers' self-reported fish consumption rates, durations, and measured PCB levels in fish from the Kalamazoo River, using the same model as used in the risk assessment. The calculations described here are performed in the spreadsheet Phase_2.wb3 (Appendix B.18).

PCBs accumulate in the body, so PCB levels show a significant correlation with age. Therefore, in order to determine whether there is a difference in the measured PCB levels between two sets of blood samples, the general increase in PCB levels with age must be taken into account. The accumulation of PCBs in the body over time may be modeled by using the metabolic and nonmetabolic excretion rates for each PCB congener, as discussed in Section 6.3.2. For this analysis, a constant intake rate of PCBs is assumed, with a congener mix corresponding to the 75% bass, 25% carp Aroclor mix discussed in Section 6.3.2 and shown in Table 6.11. The total accumulation for a unit annual dose of total PCB is shown in Figure 6.22. This figure can be applied to any constant dose rate — for example, for 1 mg/(kg body weight) PCBs per year (about 2.74 µg/kg-day in an adult), the numbers on the y-axis should be interpreted as mg/kg (total body burden/body weight); for other dose rates, the body burden is proportional to the dose rate.

[45] The documentation (MiCPHA, 2000a) states that Phase II was carried out on anglers. The identification codes and interview timing in the databases (MiCPHA, 2000b,c) suggest that some of the Phase II participants were the family of anglers, in addition to the anglers themselves. For our analysis, it makes no difference — the angler status of the participants is irrelevant; what matters is their blood PCB concentration and the fish they ate from the Kalamazoo, as determined during the Phase II interview.

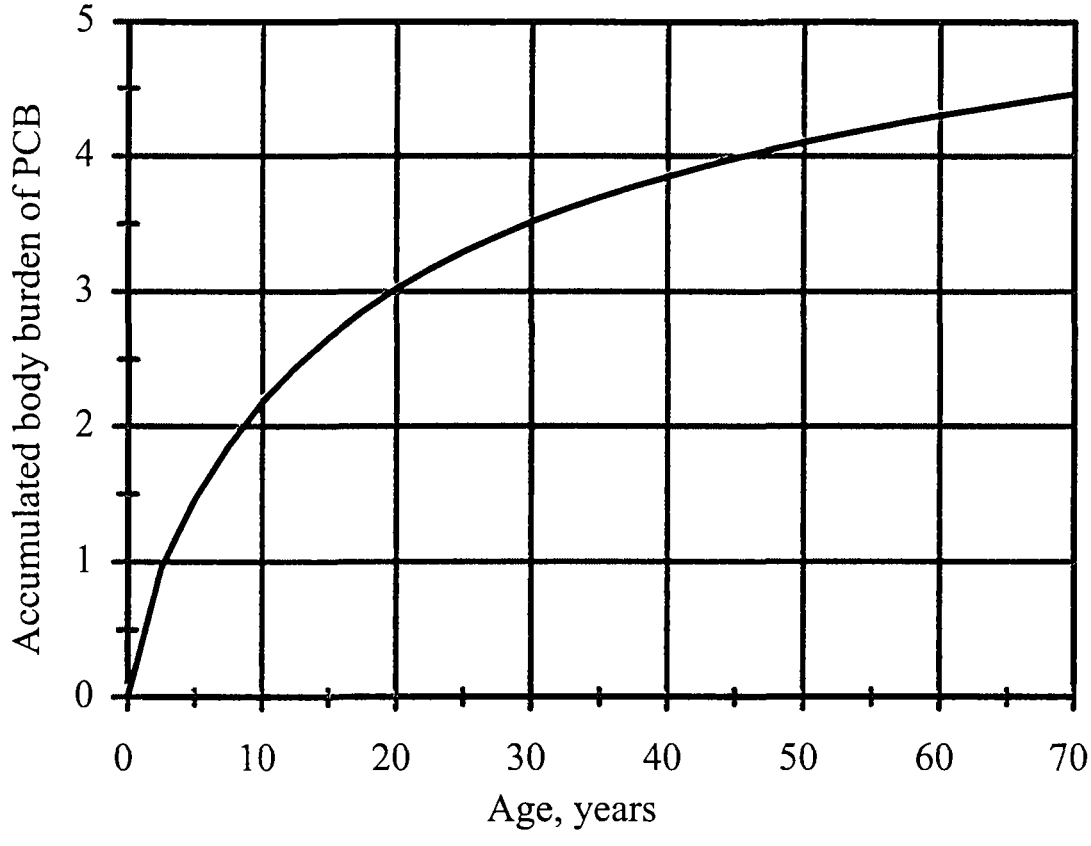

Figure 6.22 Accumulation of PCBs with age, for unit dose per year.

In order to convert a value for total accumulation of PCBs in the body to blood serum concentrations it is assumed that the PCBs in both the blood serum and the rest of the body are contained in lipids at equal lipid-based concentrations. The conversion is then simply based on the fraction of the body's lipids in the serum. Mean body lipid content and mass are taken as respectively 21.3% and 70kg for males, 32.7% and 58 kg for females (Brown, 1997); the serum lipid content is taken as 6.29 and 6.32 g/L for males and females, and the serum volume as 2.46 and 2.13 liters for males and females (Lentner, 1984). The resulting conversion factors are 19.3 and 29.5 ng/mL (ppb) serum PCB concentration per mg/kg (total PCB body burden/body weight) for males and females respectively. The values given correspond to approximately 30 to 40 year-olds. Some of these parameters values vary somewhat with age and body weight, but it is not known exactly how the overall conversion factor would vary with age or body weight. The variations are, however, much smaller than the inter-individual variability in PCB blood

concentrations. The size of these conversion factors is confirmed by direct experimental evidence in which one volunteer was administered a dose of PCBs, and subsequent blood measurements were taken (Buhler *et al.* 1998, as reported in ATSDR, 2000).

Measured blood serum concentrations of PCBs in a population are generally found to be approximately lognormally distributed, and this is true for the measured concentrations in the fish-eaters and non-fish-eaters in the Phase II Kalamazoo River study. The distribution of serum concentrations is presumably due to differences in dose rates, to differences between individuals in metabolic and excretion rates of PCBs, and to the differences in conversion factor previously discussed. For a constant dose rate, the expected blood serum concentration would increase in proportion to the curve shown in Figure 6.22. For different dose rates for different people, and differences between people, we expect the median measured serum concentration to increase in proportion to that curve, with a substantial variation between individuals. In reality, it is unlikely that the dose rate remains constant over any appreciable period, but we use this as an reasonable approximation — the dose rate will be interpreted as the average over the period of exposure.

For the fish-eaters and non-fish-eaters in the Phase II survey, we therefore initially approximate their blood concentrations by a model of the form:

$$\ln(C) = \ln\left(Af(t) + B\right) + \varepsilon \tag{6.28}$$

where C is the concentration for an individual, $f(t)$ is the curve given in Figure 6.22 multiplied by the conversion from body burden to blood concentration (so $f(t)$ is the blood concentration expected at age t for unit average dose rate up to that age), and ε is a normally distributed error term with mean zero. A and B are parameters, the former estimating the population median dose rate up to age t, the latter being an offset from the model that should be zero if the model is an adequate description.

For this analysis, we defined fish-eaters as those who indicated any consumption at any time of fish from the relevant stretch of the Kalamazoo — they said they had eaten such fish within the last twelve months, had changed their fish consumption within the last five years, indicated a first or second most frequently eaten type of such fish, indicated a length of time for which they had been eating such fish, or gave a non-zero count for the total of such fish eaten in the last twelve months. Non-eaters were those who answered the questionnaire but did not fall in any of the above categories. In either case, the interviewee had to be recorded as over the age of ten at time of interview, to have a recorded sex, and to have a valid PCB blood analysis. The age cut-off was designed to exclude some apparently contradictory responses, for example an age of 0.4 years and length of eating fish of 19 years. Finally, the single individual with 73 ppb PBCs in blood identified as probably exposed to some other source of PCBs (based on a discrepant congener distribution, MiCPHA, 2000a, page 76), has been excluded. These definitions differ substantially from those used for the analysis given in MiCPHA (2000a), which only looked at a distinction on eating fish within the last twelve months.

Applying this model (using likelihood methods) to the data from the Phase II study for the fish-eaters and non-fish-eaters separately (and combining men and women), we found that the term B was not significantly different from zero ($p>0.05$, two-sided, likelihood ratio test) as expected if $f(t)$ is an accurate accumulation curve.[46] We also found that there was no difference in the standard deviations of logarithm of concentration in fish-eaters and non-fish-eaters ($p>0.05$, two-sided, likelihood ratio test), and that the value of A differed significantly between fish-eaters and non-fish-eaters ($p=0.005$). The value of A corresponded to a dose rate (for a 70-kg adult) of 0.023 µg/kg-day for non-fish-eaters, and to 0.046 µg/kg-day for fish-eaters, a difference of 0.023 µg/kg-day.

The fish-eating anglers answered questions about how much fish they eat from the Kalamazoo River per year, what type of fish they eat, and how long they have been eating these fish. Because PCB concentrations of fish in the river have been measured, it is possible to predict the expected difference in average dose rates of PCBs between the fish-eaters and the non-fish-eaters using the risk assessment model, and incorporate a term based on fish consumption in the PCB blood concentration model just given.

Based on measurements in smallmouth bass and carp in 1993, 1997, and 1999, concentrations of PCBs in Kalamazoo River fish have been decreasing at a rate of approximately 4.81% per year (Section 6.2.4). The average concentration in fish eaten over a period from calendar time $t-\tau$ to t, given a concentration c at calendar time T is then

$$c\frac{\exp\left(-\beta(t-\tau-T)\right)-\exp\left(-\beta(t-\tau)\right)}{\beta\tau} \tag{6.29}$$

By assuming that the fish-eaters consumed fish from the Kalamazoo River at a constant rate over the entire period for which they claim to have eaten such fish, it is possible to predict the amount of fish-derived PCBs that should have accumulated in each angler's body when the blood serum sample was taken.[47]

---

[46] The p-values cited in this paragraph have to be interpreted with considerable reserve. The samples of fish-eaters and non-fish-eaters are highly self-selected, so that at least one fundamental requirement for interpreting p-values as probabilities is likely violated. This self-selection also casts doubt on the representativeness of the values obtained for the whole population of fish-eaters or non-fish-eaters. Analysis of males and females separately using this model suggests that male and female fish-eaters are very similar, but that male and female non-fish-eaters are very different (with the nine female non-fish-eaters having a median dose rate similar to the fish-eaters, but with a much smaller variance — see spreadsheet Phase_2.wb3, Appendix B.18). We have chosen to analyze males and females together, however, in the belief that this difference is more likely to arise from the way the sample was assembled.

[47] This is just the model used in the risk assessment, Sections 6.1, 6.2.5, and equation 6.7.

For this analysis, the mean fish PCB concentrations given in Tables 6.3, 6.5 and 6.7 for ABSAs 3 through 9 have been used. Each mean value was adjusted to 1995 using a decay constant of 4.81% yr^{-1}, then all available values for the type of fish were averaged. The Phase II questionnaire obtained the kind of fish (walley, sucker, carp, bass, pike, panfish, catfish, bullheads, turtles, and other; as for Phase I, see Section 6.5.2) eaten most often. The concentration of "other" fish was set to a weighted average concentration of the other fish, using the average fractions eaten as determined from the Phase I survey (see Section 6.5.2). As for Phase I, bullheads were included in the catfish category.

Based on the fish eaters self-reported consumption rates, fish PCB concentrations, and a combined PCB survival factor of 0.75 to account for cooking losses, the predicted increment in median dose rate caused by eating fish from the river and during the period of fish-eating is 0.072 µg/kg-day, while the lifetime average contribution to median dose rate is 0.015 µg/kg-day. This last value is not much different than the previously determined difference in median lifetime dose rates for the fish eating and non-fish eating groups of 0.023 µg/kg-d. Thus, the overall difference between the PCB blood serum levels of the fish eaters and non-fish eaters might be explained by fish consumption rates and PCB concentrations using the risk assessment model, if it is valid to average PCB dose rate estimates over whole lifetimes and populations.

A better test may be performed to determine whether reported fish consumption rates and durations can be used to predict the variability in PCB concentrations among the fish eaters. This may be done by modifying the equation for predicted serum PCB concentration to:

$$\ln(C) = \ln\left(Af(t) + GF + B\right) + \varepsilon \tag{6.30}$$

where F is the computed increment in PCB concentration in serum based on the reported fish consumption, concentrations of PCBs in the fish and its rate of change with time, an absorption factor of 0.75, and the accumulation model applied over the period of fish-eating. G is a parameter that should have the value unity if the modeling of accumulation of PCBs from fish is entirely accurate, and A should in this model be the same for the fish-eaters and non-fish-eaters, since the only difference between these groups should be their fish consumption — their intake of PCBs from other sources should be similar, at least on average.

Once again, the standard deviations are not significantly different between the two groups, and B is not significantly different from zero ($p>0.05$, two-sided 95% confidence). However, the value of G is also not significantly different from zero, either with A values allowed to differ between fish-eaters and non-fish-eaters, or with A values forced to be identical for the two groups. Thus while the inclusion of a modeled increment in PCB levels due to fish consumption can laragely explain the overall population difference in PCB blood serum concentrations between the fish-eating and non-fish-eating anglers, it does not significantly explain the variations in PCB concentrations among the fish-eating-anglers.

If instead of demanding that the modeled PCB blood levels match on an individual basis, but simply reproduce the median blood level of the population of non-fish-eaters and fish-eaters, then the coefficient G turns out to be approximately 0.13 (which is within the uncertainty range of this coefficient for the model where the coefficient A is forced to be the same for fish-eaters and non-fish-eaters), suggesting that the fish consumption model overestimates PCB intake by a factor of about 7. With this value for the coefficient, the model matches the variability of blood PCB concentrations for non-fish-eaters if the error term ε has a standard deviation of approximately 1.15, corresponding to a random, unexplained variation of a factor of 3.2 in blood PCB concentrations; however, there is then still an unexplained variation of a factor of about 2 in the fish-eaters.

We have not identified the explanation(s) for this failure to explain individual variation in PCB serum concentrations, or the potential overestimate of median intake of PCBs by the risk assessment model. Some plausible possibilities include, but are not limited to, some combination of:

- Additional variation that was not incorporated in this model. For example, we used an average PCB loss from cooking, rather than accounting for variability between people due to their different cooking methods.
- Some of the fish eaten by some of the Phase II participants might have come from reaches of the Kalamazoo that are less contaminated (*e.g.* Morrow Lake).
- The reported fish intakes, length of eating fish, or some combination of these values in the Phase II Kalamazoo River Angler Survey may be overestimates.
- Absorption of PCBs from fish might be substantially lower than the 100% assumed in the modeling, or the mode of exposure might affect metabolism in some way. Some such dietary effect is indicated in a study on infant rhesus and cynomolgus monkeys, where a currently unexplained difference was observed in blood PCB concentrations between animals dosed in the liquid diet or in corn oil (Arnold *et al.*, 1999).
- The modeled relation between PCB intake and PCB body burden may be incorrect. It is based on published estimates for PCB congener metabolic rates, but they may be incorrect for low PCB dose rates. The congener composition of the PCBs that we use corresponds to the mixtures of commercial Aroclors that most closely match chromatograms of PCBs measured in fish. We do not know the effect of any mismatch between the assumed and the actual congener mix.

To summarize, when the risk assessment model is applied in conjunction with a detailed model relationship between PCB intake and blood concentration, the combined modeling does not adequately explain the individual measured blood concentrations in the Phase II Kalamazoo River Angler Survey. When these two models are used to estimate the population median blood PCB concentration, the risk assessment model may overestimate PCB intake by a factor of about 7 compared with the measurements in the Phase II Kalamazoo River Angler Survey. However, the median difference in blood concentrations between fish-eaters and non-fish-eaters in this

survey can also be adequately explained by averaging the risk-assessment model estimates of fish consumption over a lifetime, as is done for lifetime cancer risk in the risk assessment itself.

The risk assessment model in conjunction with the intake/blood concentration model thus appears to be inadequate to explain individual blood PCB concentrations, and may overestimate intakes by a factor of 7 when applied using individual data. However, for lifetime population average estimates of dose, the risk assessment model may be reasonably accurate.

7 *Exposures that are de minimis*

7.1 Vapor exposure from former impoundments

Vapor evaporation from the former impoundments followed by vapor inhalation is a minor pathway of incremental exposure to PCBs, since the surface soil in the former impoundment areas has been exposed to the atmosphere long enough to reach equilibrium with background levels of PCB in the atmosphere. At least six sets of air monitoring data are available in connection with the Kalamazoo site, all performed adjacent to Portage Creek and the Kalamazoo River under effectively worst-case conditions — that is, either during midsummer and/or during extensive remediation of some of the most-contaminated areas of the site. Around the Allied Paper, Inc., Operable Unit, adjacent to Portage Creek, air monitoring every sixth day between June 6, 1993 and August 29, 1993, indicated total air concentrations of PCBs from 0.9 to 6.4 ng/m^3 (BBL, 1994a). Sampling over the same period around the Willow Boulevard/A-Site Operable Unit, adjacent to the Kalamazoo River, detected total concentrations of PCBs in air of 0.5 to 2.9 ng/m^3 (BBL, 1994b). More recently (1998–2000), air monitoring was conducted during the extensive response or removal actions at the Former Bryant Mill Pond area, the Former Georgia-Pacific Corporation Mill Lagoons, the Former Allied Paper, Inc. King Mill Lagoons, and the Willow Boulevard/A-Site OU. The few measurements above 20 ng/m^3 in these monitoring data are clearly related to the excavation and movement of highly contaminated soils, not to normal emissions from soil or the adjacent waterways. During periods with no movement of highly contaminated soils, the air PCB concentrations were similar to background levels.

Thus, at the former impoundments, the air concentration of PCBs can be expected to be well below 20 ng/m^3. Even exposure at an average of 20 ng/m^3 results in negligible risk estimates for the scenarios evaluated here (see the spreadsheet Other_exposures.wb3, Appendix B.11). For the hunter/fisher scenario, exposure for 8 hours/day for the 20 days/year on a former impoundment with an effective exposure period of 40 years leads to an average dose rate of approximately 6×10^{-8} mg/kg-day (20 m^3/day inhalation, 70 kg body weight). This corresponds to a risk estimate of less than 2.4×10^{-8}. Similarly, for the trespassing gardener (8 hours/day, 100 days/year, 25 years effective exposure), an exposure concentration of 20 ng/m^3 would lead to a lifetime inhalation dose of 1.9×10^{-7} mg/kg-d, corresponding to a lifetime risk estimate of less than 7.5×10^{-8} (upper bound potency 0.4 kg-day/mg by inhalation, IRIS, 2001, and U.S. EPA 1996).

7.2 *Vapor exposure from river water*

The HHRA (MiDEQ, 2000) raises the possibility that elevated levels of airborne PCBs might exist downwind of the dams (or remaining dam sills) along the Kalamazoo River due to the volatilization of PCBs in the river. The latest measurements (year 2000) indicate a mean concentration of PCB's in the river water of 0.023 µg/L, treating non-detects as ½ the detection limit (LTI, 2000; these calculations are performed in spreadsheet Surface_water.wb3, Appendix B.19).

Either or both of two factors may limit the amount of PCBs that could be volatilized from the river as it flows over a dam: the rate of PCBs arriving at the dams (the product of river's flow rate and the PCB concentration), and the equilibrium concentration of PCBs in the air based on the aqueous concentration and the Henry's law constant for the PCB mixture in the river. The mean flow rate of the river at the Fennville gaging station just downstream of Lake Allegan is 1450 cubic feet per second, or 41,000 liter/s (Blasland & Bouck, 1992b), yielding a total amount of PCB passing over the dam at Lake Allegan of 0.94 mg/s (and smaller amounts upstream). A composite, dimensionless Henry's law constant of 6.2×10^{-3} was derived from directly measured and estimated congener-specific Henry's law constants (Brunner *et al.*, 1990, see spreadsheet PCB_congener_data.wb3, Appendix B.13), the overall mixture of detected Aroclors from the surface water measurements performed from 3/7/2000 through 6/1/2000 (LTI, 2000), and the congener composition of each Aroclor (Frame *et al.*, 1996; ATSDR, 2000). The product of the mean concentration of PCBs in the river water of 0.023 µg/L and the Henry's law constant gives a maximum atmospheric PCB concentration of 0.14 $\mu g/m^3$ directly above the water.

To estimate the atmospheric dispersion of PCBs the dam was treated as a volume source of width 50 m (across the dam) by 4 m height by 4 m width. This volume source was in turn approximated by 26 point sources located every 2 m along the width of the dam and arranged to incorporate (separate vertical and horizontal) virtual distances in the dispersion equations corresponding to standard deviations of 2 m in the vertical and horizontal directions at the point source locations. Concentrations were modeled at receptors located radially the center of the dam at 0°, 45° and 90° to the line of the dam, at radial distances of 30 m, 100 m, 200 m, 500 m, and 1000 m, and at a short radial distance corresponding to a nearest distance of 5 m from the dam (that is, at 30 m, 7.071 m, and 5 m for the angles 0°, 45° and 90° to the line of the dam). Concentrations from each of the 26 sources for a given wind direction were computed for receptors at the same height of the source using the usual Gaussian plume formula :

$$\frac{\chi}{Q} = \frac{1}{\pi\sigma_z(x+v_v)\sigma_y(x+v_h)u}\exp\left(-y^2 \Big/ 2\sigma_y^2(x+v_h)\right) \qquad (7.1)$$

where the terms are

χ	Concentration (ML^{-3}) from a single source,
Q	Emission rate (MT^{-1}) of the single source,
y	Cross-wind distance to receptor (L),
x	Source-receptor downwind distance (L),

v_v	Vertical virtual distance, (L), adjusted so that at x=0 the vertical standard deviation is 2 m,
v_h	Horizontal virtual distance, (L), adjusted so that at x=0 the horizontal standard deviation is 2 m,
$\sigma_z(x+v_v)$	Vertical plume standard deviation (L), at downwind distance x with vertical virtual distance v_v,
$\sigma_y(x+v_h)$	Horizontal plume standard deviation (L), at downwind distance x with horizontal virtual distance v_h,

where the plume standard deviations are those used in the U.S. EPA dispersion models, and are evaluated as shown for the downwind source-receptor distance plus the virtual distances. Corresponding uniform-wind-rose long-term-average concentrations were computed for each source at each receptor using angle-averaged (cross-wind integrated) form of equation 7.1:

$$\frac{\chi_{u,a}}{Q} = \frac{1}{2\pi r \sigma_z (r + v_v) u} \sqrt{\frac{2}{\pi}} \tag{7.2}$$

where the new terms are:

$\chi_{u,a}$	uniform-wind-rose, long-term-average concentration (ML^{-3}) from a single source,
r	source-receptor distance (L) for the single source,

These concentration values were computed for each of the 26 sources across the dam, and summed to obtain an estimate for the concentration from total emissions from the water flowing over the dam. The emission rate Q was adjusted to be either 1/26 of the rate of flow of total PCBs flowing over the dam, or so that the maximum concentration at 5 meters from the dam (for any wind direction) did not exceed 0.14 $\mu g/m^3$, the limit imposed by Henry's law (calculations are performed in spreadsheet Other_exposures.wb3, Appendix B.11). These calculations were performed for individual wind speeds, and averaged over wind speeds using the wind speed distribution from one year (1992) of hourly measurements taken at Grand Rapids/Kent County International Airport, assuming that calm periods corresponded to a wind speed of 1 m/s (in practice the Henry's law constraint dominated for all wind speeds in every case, so that concentrations are independent of wind speed). All calculations were performed also for all stability classes, as though the same stability class applied continuously. As a result of the extended nature of the source, for close-in receptors the unstable classes gave higher long-term average concentration estimates, while for more distance receptors the stable classes gave higher averages.

The uniform wind rose approximation may slightly underestimate long-term average concentrations in particular directions, so the uniform wind rose average was multiplied by an additional factor of 1.47 to account for non-uniformity. This factor was computed from the wind rose data from the same data set (the hourly meteorological data for Grand Rapids/Kent County International Airport in 1992), as the maximum ratio of the fraction of time the wind was in any 10° sector to the average fraction of time in any such sector.

Taking the maximum over stability classes resulted in the values shown in Table 7.1 for long-term average air concentrations at various distances and angles from the dam.

Table 7.1 Maximum modeled total PCB concentrations, in ng/m^3, near a dam on the Kalamazoo.

distance from center of dam (m)	Angle (degrees) from line of dam		
	0	45	90
30	40	20	17
100	2.2	3.4	3.8
200	0.6	1.2	1.3
500	0.13	0.27	0.31
1000	0.043	0.088	0.099

The values for 30 m might be applicable for estimating exposures of persons fishing or otherwise engaged in recreation near the dams. Based on USGS maps, the nearest permanent buildings to any dam correspond to the 100 m distance, while a distance of 200 m encompasses several permanent structures at several of the dams.

The highest lifetime average daily dose due to inhalation of the above levels of PCBs can be calculated for a lifetime exposure at 100 m from the source at a location 45 degrees from the line of the dam (it is assumed that no one lives in the center of the river). Multiplying the average concentration of 3.4 ng/m^3 by a nominal inhalation rate of 20 m^3/day and dividing by a nominal body weight of 70 kg gives an average daily dose of 9.8×10^{-7} mg/kg-day, corresponding to a lifetime risk estimate of less than 3.9×10^{-7} (using the upper-bound potency for inhaled PCBs of 0.4 kg-day/mg; IRIS 2001 and U.S. EPA 1996). A person exposed at the bank of the river, 5 m from the edge of the dam where the concentration is 40 ng/m^3, breathing 10 m^3 during an active 8 hours per day, 50 days per year, for an effective exposure period of 30 years, would have a lifetime average daily dose of less than about 3.4×10^{-7} mg/kg-day, corresponding to a lifetime risk of less than 1.3×10^{-7}. Such risk estimates are sufficiently low to be considered negligible.

7.3 *Exposures during swimming*

The dose rate (d) for PCBs absorbed through the skin while swimming can be estimated using a simple steady state model as:

$$d = CK_p A \tag{7.3}$$

where K_p is the permeability coefficient for PCBs in water through the skin, C is the concentration of PCBs in the water, and A is the exposed skin area. The permeability coefficient for a PCB congener can be estimated in cm/hr from the correlation:

$$\log K_p = -2.72 + 0.71 \log K_{ow} - 0.0061 M_w \tag{7.4}$$

where K_{ow} is the octanol–water partition coefficient, and M_w is the molecular weight (g/mol) of the congener (U.S. EPA, 1992a). Values for congener specific K_{ow} were taken from Hawker and Connell (1988). An overall K_p for PCBs in the Kalamazoo River of 0.61 cm/hr was calculated (see spreadsheet PCB_congener_data.wb3, Appendix B.13) based on the overall mixture of detected Aroclors from the surface water measurements performed by LTI (2000) from 3/7/00 through 6/1/00, and the congener composition of each Aroclor (ATSDR, 1998).

The mean PCB concentration measured in waters of the Kalamazoo River is 0.023 µg/L based on LTI (2000). This value includes both soluble PCBs and those associated with suspended solids, the latter of which are not likely to be absorbed as readily through the skin. Assuming a swimmer having a body surface area of 23,000 cm^2 (the 95th percentile adult male value, U.S. EPA, 1992a), 100% immersed in the water containing 0.023 µg/L PCBs yields an absorbed PCB dose rate of 0.32 µg/hr (see spreadsheet Other_exposures.wb3, Appendix B.11). Applying the central recommended defaults (U.S. EPA, 1992a) for an adult's swimming event time (0.5 hr), frequency (1 event/day, 5 events/year), and duration (9 years) gives a total lifetime absorbed dose of 7.2 µg, corresponding to a dose rate during exposure for a 70 kg adult of 3.1×10^{-5} µg/kg-day, far below the health protective value of 0.05 µg/kg-day. The lifetime average dose rate (70 year lifetime) is 4.0×10^{-9} mg/kg-day, corresponding to a risk of less than 8.0×10^{-9} (using the upper bound potency estimate of 2 kg-day/mg, IRIS 2001 and U.S. EPA 1996). It is highly unlikely that sub-populations characterized by the upper defaults for swimming time, frequency, and duration (*e.g.* competitive swimmers) would exclusively use the Kalamazoo River for swimming. However, even applying the recommended default exposure parameters for such swimmers (1 hr/event, 1 event /day, 150 days/year, 30 years duration) gives a dose rate during exposure of 0.0019 µg/kg-day, still well below the health protective value of 0.05 µg/kg-day. In this case, the lifetime dose rate would be 8×10^{-7} mg/kg-day, corresponding to a risk of less than 1.6×10^{-6} (using the upper bound potency estimate of 2 kg-day/mg, IRIS 2001 and U.S. EPA 1996). These risk estimates are sufficiently low to be considered negligible.

8 *Comparison with Michigan's screening-level HHRA*

8.1 Introduction

This chapter examines the similarities and differences between this assessment and the screening-level assessment produced by the Michigan Department of Environmental Quality ("Final Human Health Risk Assessment" for the Allied Paper, Inc./Portage Creek/Kalamazoo River Superfund Site (HHRA; MiDEQ, 2000). We here point out the sources of such differences and their consequences. Our detailed assessment suggests that the plausible high-end risks associated with this site range from negligible to none. The MiDEQ screening-level assessment suggests instead that site risks should be reduced. However, flaws in the MiDEQ HHRA render its risk estimates unrepresentative, unreliable, or both.

8.2 Anglers who eat the fish they catch

8.2.1 Results of this assessment

The major exposure to PCBs from the Kalamazoo river site is to anglers who eat the fish they catch. We have performed a probabilistic analysis of this exposure route and report the variation in risks among the population of such anglers, and the uncertainties in those risks. Our evaluation is based on the information (what fish they eat, how often, for how long, and so forth) that was provided by those anglers in the Kalamazoo River Angler Survey and Biological Testing Study (MiCPHA, 2000a). We evaluated lifetime average dose rates of PCBs, and dose rates of PCBs during the period that each angler actually eats fish from the Kalamazoo. These dose rates were used for comparison with fixed cancer potency estimates and safe doses respectively, to evaluate lifetime risk estimates and hazard indexes. In addition, we evaluated the variability and uncertainties in the cancer potency estimates and the safe doses, in order to obtain a more realistic estimate of the true variabilities and uncertainties of lifetime risk estimates and hazard indexes.

In the detailed part of this assessment we separately evaluated variabilities and uncertainties for both doses and risks. We also combined them to generate an uncertainty distribution for doses and risks for a randomly chosen individual. For comparison with the HHRA, we concentrate on the estimates for a random individual, since that corresponds best with the approach taken in the HHRA (where no distinction was made between variability and uncertainty).

For the random individual, the uncertainty distribution for lifetime average dose rates (Figure 6.15) shows the median (50th percentile) estimate to be 0.0028 μg/kg-day, and the 90th percentile estimate (a plausible, high-end estimate) to be 0.041 μg/kg-day. These are both smaller than a dose rate of 0.05 μg/kg-day, the value endorsed as safe for long-term exposure by the Michigan Environmental Science Board (Fischer *et al.*, 1998). The dose rate during exposure (as opposed to averaged over a lifetime) can be considerably higher — the uncertainty distribution for the random individual (Figure 6.16) has median 0.052 μg/kg-day, and a 90th percentile at 0.31 μg/kg-day — although such dose rates are for shorter periods.

A full analysis requires incorporating the variability and uncertainty in the toxicity values with which doses are compared. In this accounting for carcinogenic potency (Section 4.2), we included the different carcinogenic potencies of the different Aroclors (conditional on PCBs being human carcinogens at all at these doses), the within-species variability, and the cross-species uncertainty. For non-cancer toxicity (Section 4.3), we incorporated the variation with exposure time, the variability of individuals, the uncertainty in interspecies extrapolation, and the uncertainties of extrapolation from LOAEL to NOAEL.

As for doses, for comparison of risk estimates with the HHRA the uncertainty distribution for a randomly chosen fish-eating angler (Figure 6.18) is most appropriate. The potential cancer risk ranges from completely negligible values (far smaller than 1 in 100,000,000), through a 50th percentile (median) estimate of 2.2 in 10,000,000, to a plausible, high-end, 90th percentile value of 1.7 in 100,000. For health risks other than cancer (evaluated by means of a hazard index), for a randomly chosen individual in the fish-eating angler population (Figure 6.20), the median (50th percentile) hazard index is 0.020, and the 90th percentile (plausible high-end) hazard index is 0.81. Both U.S. EPA and the State of Michigan consider hazard indices smaller than 1.0 to be acceptably small.

Finally for this exposure scenario, we examined the possible total effects on cancer rates in the entire population from current and all future exposures to the site. No direct comparison with the HHRA is possible, since the HHRA made no attempt to estimate population effects. With all variability and uncertainties accounted for, the calculations show (Figure 6.21) that the upper 90th percentile estimate for the total number of site-related cancers that might ever occur is about 2.2, with an expected value of 1.574%. Thus, it is highly likely that no cancer will ever occur in the whole population due to consumption of PCBs in fish from the Kalamazoo.

8.2.2 Comparison with the HHRA

While the risk-assessment models used in the HHRA (MiDEQ, 2000) and in our assessment are very similar, there is a large contrast in approach and implementation. As it turns out, these differences lead to marked differences in results. While we here perform a full uncertainty and variability analysis for anglers who eat fish from the Kalamazoo, the HHRA performed a screening level assessment. A comparison of some the HHRA input assumptions with the evaluation given here is shown in Table 8.1. The principal differences are in the scenarios evaluated — the HHRA constructs hypothetical "Sports Anglers" and "Subsistence Anglers," whereas we use the measured information from the State-sponsored investigation of the angler population. The HHRA assumptions as to fish meal frequency are high on the observed distributions, as is the HHRA-assumed exposure period. While use of the average PCB concentration in the HHRA is appropriate (since fish-eating anglers will be exposed to an average of many fish), the use of the single maximum concentration in such an evaluation is not. That maximum concentration in any individual fish is larger than the expected average concentration to which anyone would be exposed, and so represents an extreme on the uncertainty distribution for average concentration.

For the lifetime, excess cancer risk calculations, the HHRA used an upper bound, high risk, persistent substance cancer slope factor equal to 2.0 $(mg/kg\text{-}day)^{-1}$ — as was used here when the toxicity uncertainties and variabilities were not included. For the full probabilistic analysis, we incorporate the variation of potency for different Aroclors, the population variability observed for PCBs, and the interspecies uncertainties observed for carcinogens.

In the HHRA, the chronic, oral reference doses (RfDs) for non-cancer effects used in the hazard quotient calculations were 0.02 μg/kg-day for immunological endpoints and 0.07 μg/kg-day for reproductive endpoints. These RfDs were derived from animal exposure studies; the former for Aroclor 1254, the latter for Aroclor 1016. In contrast, for the non-probabilistic part of this document, we adopt the estimate of 0.05 μg/kg-day applicable to environmental mixtures of PCBs and endorsed by the Michigan Environmental Science Board (MESB) (Fischer *et al.*, 1998). This value is also used for comparison purposes in examining the doses estimated in the probabilistic analysis of the fish ingestion route. For the full probabilistic analysis we incorporated a variation with exposure time based on body burden, the variability of individuals, the uncertainty in interspecies extrapolation, and the uncertainties of extrapolation from LOAEL to NOAEL.

Table 8.1 Some inputs used for the risk assessment of fish-eating anglers in the HHRA (MiDEQ, 2000) and this assessment.	
HHRA (MiDEQ, 2000)	This assessment
Central Sport Anglers: 24 meals/yr High Sport Anglers: 62.5 meals/yr from the Kalamazoo Subsistence Anglers: 179 meals/yr	Observed, survey-based distribution of fish meals/year (median = 8 meals/yr; maximum = 325 meals/yr.) Extrapolated from 1 meal/yr to 1095 meals/yr by correlating with the length of time eating fish. 24 meals/yr: 74th percentile of observed 62.5 meals/yr: 93rd percentile of observed. 179 meals/yr: 97th percentile of observed.
Exposure duration 30 years, with assumed additional duration of 9 years due to body burden.	Duration distribution estimated from survey results. Additional duration computed based on cumulative body burden. 30 years is at the 96th percentile from survey. The maximum additional duration is about 7.6 years, based on cumulative body burden of PCBs.
Fish eaten: Smallmouth Bass; or 75% smallmouth bass and 25% carp	Fish reported by anglers in the survey (classified as walleye, sucker, carp, bass, pike, panfish, catfish, or turtle). Individual reports used to capture variability.
Average concentrations, or maximum concentration found in a single fish, by individual ABSA,[a] 1993 and 1996 data only.	Average concentration by ABSA, with its uncertainty distribution. Data from 1993, 1996, and 1999.
Concentrations assumed constant	Observed time trend, with its uncertainty distribution.
PCB survival through cooking, 78%	PCB survival through cooking variable, from 100% to 26%, based on measurements encoded as distributions. Overall mean 73%.
Meal size 0.225kg	Distribution of measured values. Mean 0.246 kg.

[a] Aquatic Biota Sampling Area

The results obtained in the HHRA are summarized in Table 8.2. Since the exposure duration used as an input to the HHRA is at the 96th percentile of the variability distribution for anglers, and other inputs are also high on their respective distributions, it is not surprising that the values obtained in the HHRA are extremely high on the variability/uncertainty distributions shown in Figures 6.17 through 6.20. For example, for excess lifetime cancer risk, the values estimated by the HHRA for "average fish concentration" and "subsistence angler" would be expected to be at about the 99.9th percentile, given that the meal frequency and exposure period used for this scenario are at the 96th and 97th percentiles of their observed distributions, respectively.[1] For a random individual the distributional analysis performed here shows the upper end result obtained in the HHRA for this combination (4.5×10^{-3}) to be above the 99.8th percentile, if we include all the uncertainties of the potency estimate. If we omit the uncertainties of the potency estimate, as is done in the HHRA, the HHRA upper end result for "average fish concentration" and "subsistence angler" is above the 99.9th percentile.

Even the HHRA results for the "average fish concentration" and "Central Tendency Sport Angler" turn out to be extreme. For example, the upper end of the risk range obtained for this combination in the HHRA is 7.9×10^{-4}. On the full distribution of uncertainties for a random angler this is above the 99.2th percentile. Alternatively, if we omit the uncertainties of the potency estimate, to correspond more closely with the HHRA, it is above the 99.3th percentile. Clearly, the results of the HHRA are at extreme upper-ends of the distribution, and far above typical definitions of plausible high-end estimates used for site management decision-making.

The HHRA makes no attempt to estimate possible total effects on cancer in the entire population from current and all future exposures to the site. As such, it fails to provide decision-makers with potentially important information.

[1] The probability for both meal frequency and exposure period to exceed the HHRA values would be approximately $0.04 \times 0.03 = 0.0012$ if those two parameters were independent, so the result would be at about the $100 \times (1-0.0012) = 99.88$ percentile.

Table 8.2 Summary of results from the HHRA (MiDEQ, 2000).				
		Central Tendency Sport Angler	High End Sport Angler	Subsistence Angler
Excess Lifetime Cancer Risk	average fish concentration	2.3×10^{-4} to 7.9×10^{-4}	4.6×10^{-4} to 1.6×10^{-3}	1.0×10^{-3} to 4.5×10^{-3}
	maximum fish concentration	5.8×10^{-4} to 1.8×10^{-3}	1.2×10^{-3} to 3.7×10^{-3}	3.3×10^{-3} to 1.0×10^{-2}
Hazard Quotient – Immunological Endpoint	average fish concentration	8.1 to 36	16 to 72	46 to 200
	maximum fish concentration	35 to 81	53 to 160	150 to 460
Hazard Quotient – Reproductive Endpoint	average fish concentration	2.3 to 10	4.7 to 21	13 to 58
	maximum fish concentration	9.9 to 23	15 to 47	42 to 160

8.3 Other scenarios

8.3.1 Results of this assessment

In addition to the eating of fish by anglers, this assessment also examined exposure to PCBs present in the soils of the former impoundments by using screening level assessments (Section 5) for hunter/fishers and trespassing gardeners, the expected most highly exposed individuals. For three further potential exposures (exposure to vapors from river water, vapors from the impoundments, and exposure through swimming), screening level assessments (Section 7) showed negligible exposures and risks.

For the hunter/fisher scenario, exposures through soil contact and ingestion were considered (dust inhalation was dismissed as negligible). Using conservative parameter values (Section 5.2) and a soil concentration of PCBs of 36.0 mg/kg (the upper 95th percentile confidence estimate for the average soil concentration at the Plainwell former impoundment), the average dose rate during the period of exposure is 0.0024 µg/kg-day, substantially lower than the health-protective value of 0.05 µg/kg-day. The lifetime cancer risk estimate is 2.8×10^{-6}, well within the range of acceptable values. For the other two impoundments, UCL95 estimates of the average soil concentrations of PCBs for Otsego and Trowbridge are 21.9, and 29.3 mg/kg respectively, leading to lower exposure and risk estimates.

With conservative assumptions for exposure, the trespassing gardener (Section 5.3) was found to have an average daily dose rate during exposure of 0.15 μg/kg-day, approximately three times the protective value of 0.05 μg/kg-d. Taking account of the effective exposure period, the upper bound lifetime risk estimate is 10×10^{-5}, within the acceptable range of values for the U.S. EPA, but higher by a factor of about 10 than Michigan's limit for waste sites.

8.3.2 Comparison with the HHRA

The HHRA (MiDEQ, 2000) examines the potential effect on nearby residents and recreationalists who are exposed by ingestion, inhalation of dust, and dermal absorption of PCBs from contaminated soils by using screening-level scenario analyses. Exposure assumptions for these scenarios are summarized in HHRA Tables 3-4, and 3-5 respectively,[2] and the HHRA results are summarized here in Table 8.3. The modeling methods and parameters involved are similar to those used in our screening-level assessments. However, the exposure point concentrations used (MiDEQ, 2000, Section 3.5.3, page 3-19) were the average and maximum concentrations presented in Tables 2-1 and 2-3 of the HHRA (MiDEQ, 2000, page 2-10).

Unfortunately, the HHRA has not distinguished in its scenarios between the former impoundments and the other floodplain soils outside those impoundments. Since the impoundments are State-owned lands, it is highly unlikely that a residential scenario, or even the particular recreational scenario, envisioned by the HHRA applies to this land. Moreover, in evaluating the exposure point concentrations, no account was taken of the differences between the areas within the former impoundments and those outside them, leading to incorrect statistical treatment of the data in evaluating exposure point concentrations. Section 5.1 gives some indication of the complexity necessary in estimating exposure point concentrations in such circumstances.

The exposure point concentrations used by the HHRA thus do not match the scenarios evaluated. For the residential scenario, the HHRA used "average sediment concentrations" ranging from 8.4 to 10.9 mg/kg; but such average concentrations only occur within the former impoundments (where there are no residences), not outside. Even the "recreationalist" scenario, which assumes 128 days/year exposure to the soil for 24 years, is unlikely within the former impoundments in the absence of residences within them. The "maximum concentrations" used in the HHRA range from 36 (Otsego) to 85 (Plainwell) mg/kg; but examination of Section 5.1 shows that these concentrations are incorrect, in that they are well above even upper limit estimates for the average concentrations within their respective former impoundments — and it is the average concentrations that are relevant to the exposures within the former impoundments.

---

[2] Not all of the assumptions used for the calculations in HHRA Appendix A for residents are shown in HHRA Table 3-4, and the relevance of the two or three different assumptions listed in the Table 3-5 for recreationalists are not explained.

Table 8.3 HHRA (MiDEQ, 2000) results for nearby resident and recreationalist scenarios.			
		Nearby Resident	Recreationalist
Excess Lifetime Cancer Risk	average sediment concentration	3.7×10^{-5} to 5.4×10^{-5}	5.0×10^{-6} to 7.3×10^{-6}
	maximum sediment concentration	1.6×10^{-4} to 3.8×10^{-4}	2.1×10^{-5} to 5.0×10^{-5}
Hazard Quotient – Immunological Endpoint	average sediment concentration	2.0 to 2.9	0.21 to 0.31
	maximum sediment concentration	8.5 to 20	0.90 to 2.1
Hazard Quotient – Reproductive Endpoint	average sediment concentration	0.14 to 0.21	0.016 to 0.023
	maximum sediment concentration	0.61 to 1.5	0.068 to 0.16

The screening values obtained in the HHRA for these scenarios are thus substantial overestimates, since even the estimates of "average sediment concentration" do not correspond to any residential setting, while the "maximum sediment concentrations" do not correspond to exposure point concentrations anywhere. Evaluation of a residential setting requires using exposure point concentrations estimated from measurements outside the former impoundment areas. Evaluation of a "recreational" scenario within the former impoundments requires estimating average sediment concentrations within the impoundments, and selection of a scenario that corresponds to the recreations that would occur in the former impoundments.

Given these technical flaws, the HHRA estimates of risk for "nearby residents" and "recreationalists" are not reliable.

9 *References*

Abdel-Rahman, M.S., and Kadry, A.M. (1995). Studies on the use of uncertainty factors in deriving RfDs. *Hum. Ecolog. Risk Assess. 1*:614–624.

Anderson, E.L., and the Carcinogen Assessment Group of the U.S. Environmental Protection Agency (1983). Quantitative approaches in use to assess cancer risk. *Risk Analysis 3*, 277–295.

Armbruster, G., Gerow, K.G., Gutenmann, W.H., Littman, C.B., and Lisk, D.J. (1987). The effects of several methods of fish preparation on residues of polychlorinated biphenyls and sensory characteristics in Striped Bass. *J. Food Safety 8*:235–243.

Armbruster, G., Gall, K.L , Gutenmann, W.H., and Lisk, D.J. (1989). Effects of trimming and cooking by several methods on polychlorinated biphenyls (PCB) residues in Bluefish. *J. Food Safety 9*:235–244.

Arnold, D.L., Bryce, F., Stapley, R., McGuire, P.F., Burns, D., Tanner, J.R., Karpinski, K. (1993a). Toxicological consequences of Aroclor 1254 ingestion by female rhesus (Macaca mulatta) monkeys. Part 1A. Prebreeding phase: clinical health findings. *Food. Chem. Toxicol. 31(11)*:799–810.

Arnold, D.L., Bryce, F., Karpinski, K., Mes, J., Fernie, S., Tryphonas, H., Truelove, J., McGuire, P.F., Burns, D., Tanner, J.R., Stapley, R., Zawidzka, Z.Z., and Basford, D. (1993b). Toxicological consequences of Aroclor 1254 ingestion by female rhesus (Macaca mulatta) monkeys. Part 1B. Prebreeding phase: clinical and analytical laboratory findings. F*ood. Chem. Toxicol. 31(11)*:811–24.

Arnold, D.L., Bryce, F., McGuire, P.F., Stapley, R., Tanner, J.R., Wrenshall, E., Mes, J., Fernie, S., Tryphonas, H., Hayward, S., and Malcolm, S. (1995). Toxicological consequences of Aroclor 1254 ingestion by female rhesus (Macaca mulatta) monkeys. Part 2. Reproduction and infant findings. *Food. Chem. Toxicol. 33(6)*:457—474.

Arnold, D.L., Nera, E.A., Stapley, R., Tolnai, G., Claman, P., Hayward, S., Tryphonas, H., and Bryce, F. (1996). Prevalence of endometriosis in rhesus (Macaca mulatta) monkeys ingesting PCB (Aroclor 1254): review and evaluation. *Fundam. Appl. Toxicol. 31(1)*:42–55.

Arnold, D.L., Nera, E.A., Stapley, R., Bryce, F., Fernie, S., Tolnai, G., Miller, D., Hayward, S., Campbell, J.S., and Greer, I.(1997). Toxicological consequences of Aroclor 1254 ingestion by female rhesus (Macaca mulatta) monkeys and their nursing infants. Part 3: post-reproduction and pathological findings. *Food. Chem. Toxicol. 35(12)*:1191–1207.

Arnold, D.L., Bryce, F., Mes, J., Tryphonas, H., Hayward, S., and Malcolm, S. (1999). Toxicological consequences of feeding PCB congeners to infant rhesus (Macaca mulatta) and cynomolgus (Macaca fascicularis) monkeys. *Food. Chem. Toxicol. 37(2-3)*:153–167.

Atkin, C. (1994). *A survey study of anglers residing near the Kalamazoo River basin.* Michigan State University.

ATSDR (2000). Agency for Toxic Substances and Disease Registry *Toxicological profile for Polychlorinated Biphenyls (Update: November 2000).* U.S. Dept. of Health and Human Services, Public Health Service, Agency for Toxic Substances and Disease Registry. November, 2000.

Bahn, A.K., Rosenwaike, I., Herrmann, N., *et al.* (1976). Melanoma after exposure to PCBs. *NEJM 295(8)*:450.

Barnes, D.G., and Dourson, M. (1988). Reference Dose (RfD): Description and use in Health Risk Assessments. *Regul. Toxicol. Pharmacol. 8*:471–486.

Barsotti, D. and van Miller, J. (1984). Accumulation of a commercial polychlorinated biphenyl mixture (Aroclor 1016) in adult rhesus monkeys and their nursing infants. *Toxicol. 30*:31-44.

BBL (1994a). Blasland Bouck & Lee, Inc. *Allied Paper, Inc./Portage Creek/Kalamazoo River/Superfund Site Remedial Investigation Feasibility Study. Technical Memorandum 4. Allied Paper, Inc. Operable Unit Results of the Air Investigation.* May 1994.

BBL (1994b). Blasland Bouck & Lee, Inc. *Allied Paper, Inc./Portage Creek/Kalamazoo River/Superfund Site Remedial Investigation Feasibility Study. Technical Memorandum 5. Willow Boulevard/A-Site Operable Unit Results of the Air Investigation.* May 1994.

BBL (1994c). Blasland Bouck & Lee, Inc. *Allied Paper, Inc./Portage Creek/Kalamazoo River/Superfund Site Remedial Investigation Feasibility Study. Technical Memorandum 3. Results of the floodplain soils investigation.* February, 1994.

BBL (1994d). Blasland Bouck & Lee, Inc. *Allied Paper, Inc./Portage Creek/Kalamazoo River/Superfund Site Remedial Investigation Feasibility Study. Draft Technical*

Memorandum 12. Former impoundment sediment and geochronologic dating investigation. 1994.

BBL (1995). Blasland, Bouck & Lee, Inc. *Allied Paper, Inc./Portage Creek/Kalamazoo River/Superfund Site Remedial Investigation Feasibility Study. Technical Memorandum 9 — Willow Boulevard/A-Site Operable Unit.* February 1995 (Revised and Submitted April 1995).

BBL (1997). Blasland, Bouck & Lee, Inc. *Allied Paper, Inc./Portage Creek/Kalamazoo River/Superfund Site Remedial Investigation Feasibility Study. Technical Memorandum 7 — Allied Paper, Inc. Operable Unit.* August 1997.

BBL (2000a). Blasland, Bouck & Lee, Inc. A*llied Paper, Inc./Portage Creek/Kalamazoo River Superfund Site RI/FS Remedial Investigation Report - Phase I.* Draft for State and Federal Review. October 2000.

BBL (2000b). Blasland, Bouck & Lee, Inc. *Allied Paper, Inc./Portage Creek/Kalamazoo River Superfund Site RI/FS Feasibility Study Report - Phase I.* Draft for State and Federal Review. October 2000.

BBL (2000c). Blasland, Bouck & Lee, Inc. *Allied Paper, Inc./Portage Creek/Kalamazoo River Superfund Site RI/FS Supplement to the Kalamazoo River RI/FS - Phase I.* Draft for State and Federal Review. October 2000.

Bertazzi, P.A., Ribaldi, L., Pesatori, A., *et al.* (1987). Cancer mortality of capacitor manufacturing workers. *Am. J. Ind. Med. 11*:165-176.

Blasland & Bouck Engineers, P.C. (1992a). *Results of Soil Tests — Residential Lots*, letter from Mark Brown to Scott Cornelius, MDNR, March 26, 1992.

Blasland & Bouck Engineers, P.C. (1992b). *Allied Paper, Inc./Portage Creek/Kalamazoo River Superfund Site. Description of the Current Situation.* July 1992.

Brown, D.P. (1987). Mortality of workers exposed to polychlorinated biphenyls — an update. *Arch. Environ. Health. 42(6)*:333-339.

Brown, J.F. Jr. (1994). Determination of PCB metabolic, excretion, and accumulation rates for use as indicators of biological response and relative risk. *Environ. Sci. Technol. 28*:2295–2305.

Brown, R.P., Delp, M.D., Lindstedt, S.L., Rhomberg, L.R., and Beliles, R.P. (1997). Physiological Parameter Values for Physiologically Based Pharmacokinetic Models. *Toxicol. Indust. Health 13(4)*:407–484.

Brunner, M.J., Sullivan, T.M., Singer, A.W., Ryan, M.J., Toft, II, J.D., Menton, R.S., Graves, S.W., and Peters, A.C. (1996). *An assessment of the chronic toxicity and oncogenicity of Aroclor-1016, Aroclor-1242, Aroclor-1254, and Aroclor-1260 administered in diet to rats*. Columbus, OH: Battelle Study No. SC920192., Chronic toxicity and oncogenicity report.

Brunner, S., Hornung, E., Santl, H., Wolff, E., Piringer, O.G., Altschuh, J., and Brüggemann, R. (1990). Henry's law constants for polychlorinated biphenyls: experimental determination and structure–property relationships. *Environ. Sci. Technol. 24*:1751–1754.

Buck, G.M., Sever, L.E., Mendola, P., *et al.* (1997). Consumption of contaminated sport fish from Lake Ontario and time-to-pregnancy. *Am. J. Epidemiol. 146*:949-954.

Buhler, F., Schmid, P., and Schlatter, C.H. (1998. Kinetics of PCB elimination in man. *Chemosphere 17*:1717–1726.

Calabrese, E.J., and Baldwin, L.A. (1995). A toxicological basis to derive generic interspecies uncertainty factors for application in human and ecological risk assessment. *Hum. Ecolog. Risk Assess. 1*:555–564.

Calabrese, E.J., Stanek, E.J., Gilbert, C.E., and Barnes, R.M. (1990). Preliminary adult soil ingestion estimates: results of a pilot study. *Regul. Toxicol. Pharmacol. 12(1)*:88–95.

CDM, 2000. Camp Dresser & McKee. Letter *Re: Residential Sample Results* to Mr. Brian von Gunten, Environmental Response Division, Michigan Department of Environmental Quality. September 21, 2000.

Cichy, R.F., Zabik, M.E., and Weaver, C.M. (1979). Polychlorinated biphenyl reduction in Lake Trout by irradiation and broiling. *Bull. Environm. Contam. Toxicol. 22*:807–812.

Cogliano, J. (1996) *PCBs: Cancer Dose-Response Assessment and Application to Environmental Mixtures*. National Center for Environmental Assessment, Office of Research and Development, U.S. Environmental Protection Agency. Washington, DC. EPA/600/P–96/001F September 1996.

Crouch, E.A.C. (1996). Uncertainty distributions for cancer potency factors: laboratory animal carcinogenicity bioassays and interspecies extrapolation. *Hum. Ecolog. Risk Assess. 2*:103–129.

Cunnane, C. (1978). Unbiased plotting positions — a review. *J. Hydrology 37*:205–222.

Dar, E., Kanarek, M.S., Anderson, H.A., et al. (1992). Fish consumption and reproductive outcomes in Green Bay, Wisconsin. *Environ. Res. 59*:189–201.

Devroye, L. (1986). *Non-uniform random variate generation*. Springer-Verlag, NY. (ISBN 0-387-96305-7 and 3-540-96305-7). Errata and addenda available at http://jeff.cs.mcgill.ca/~luc/ at 10/18/2000.

Dorgan, J., Brock, J., Rotman, N., *et al.* (1999). Serum organochlorine pesticides and PCBs and breast cancer risk: results from a prospective analysis (USA). *Cancer Causes Control 10*:1-11.

Dourson, M.L., and Stara, M.F. (1983). Regulatory history and experimental support of uncertainty (safety) factors. *Regul. Toxicol. Pharmacol. 3*:224–238.

Emmett, E.A., Maroni, M., Schmith, J.M., *et al.* (1988a). Studies of transformer workers exposed to PCBs: I. Study design, PCB concentrations, questionnaire, and clinical examination results. *Am. J. Ind. Med. 13*:415-427.

Emmett, E.A., Maroni, M., Jeffrys, J., *et al.* (1988b). Studies of transformer workers exposed to PCBs: II. Results of clinical laboratory investigations. *Am. J. Ind. Med. 14*:47-62.

Fein, G., Jacobson, J., Jacobson, S., *et al.* (1984). Prenatal exposure to polychlorinated biphenyls: effects on birth size and gestational age. *J. Pediatr. 105(2)*:315-320.

Fischer, L.J., Bolger, P.M., Carlson, G.P., Jacobson, J.L., Knuth, B.A., Radike, M.J., Roberts, M.A., Thomas, P.T., Wallace, K.B., and Harrison, K.G. (1995). *Critical Review of a Proposed Uniform Great Lakes Fish Advisory Protocol, September 1995*. Michigan Environmental Science Board, Lansing. xii + 62p. MESB-FP3 9/9/97

Fischer, L.J., Bolger,P.M., Jacobson, J.L., Premo, B.J., van Ravenswaay, E.O. and Harrison, K.G. (1998). *Evaluation of Michigan's Proposed 1998 Fish Advisory Program, January 1998*. Michigan Environmental Science Board, Lansing. vi + 67p.

Fries G.F., Marrow, G.S., and Somich, C.J. (1989). Oral bioavailability of aged polychlorinated biphenyl residues contained in soil. *Bull. Environ. Contam. Toxicol. 43*:683–690.

Frame, G. M., Cochran, J. W., and Boewadt, S.S., 1996. Complete PCB congener distributions for 17 Aroclor mixtures determined by 3 HRGC systems optimized for comprehensive, quantitative, congener-specific analysis. J. High Resol. Chromatogr., 19:657-668. Data available from http://www.epa.gov/toxteam/pcbid/down.htm.

Gaylor, D.W., and Kodell, R.L. (2000). Percentiles of the product of uncertainty factors for establishing probabilistic Reference Doses. *Risk Analysis 20*:245–250.

Gilbert, R.O. (1987). Statistical methods for environmental pollution monitoring. Van Nostrand Reinhold, N.Y., N.Y. ISBN 0-442-23050-8.

Gladen, B.C. and Rogan, W.J. (1991). Effects of perinatal polychlorinated biphenyls and dichlorodiphenyl dichloroethene on later development. *J. Pediatr. 119*:58-63.

GLSFATF. 1993. *Protocol for a Uniform Great Lakes Sport Fish Consumption Advisory, September 1993 (Draft)*. Great Lakes States Fish Advisory Task Force, Council of Great Lakes Governors, Chicago, Illinois. MESB-FP3 9/9/97

Gold, L.S. and Zeiger, E. (1997). *Handbook of Carcinogenic Potency and Genotoxicity Databases*, CRC Press.

Gustavsson, P. and Hogstedt, C. (1997). A cohort study of Swedish capacitor manufacturing workers exposed to polychlorinated biphenyls (PCBs). *Am. J. Ind. Med. 32*:234-239.

Hawker, D.W., and Connell, D.W. (1988). Octanol–Water Partition Coefficients of Polychlorinated Biphenyl Congeners. *Environ. Sci. Technol. 22*: 382–387.

Helzlsouer, K., Alberg, A., Huang, H., *et al.* (1999). Serum concentrations of organochlorine compounds and the subsequent development of breast cancer. *Cancer Epidemiol. Biomarkers Prev. 8*:525-532.

Holmes, K.K., Shirai, J.H., Richter, K.Y., and Kissel, J.C. (1999). Field measurement of dermal soil loadings in occupational and recreational activities. *Environmental Research 80(A)*:148–157.

Huisman, M., Koopman-Esseboom, C., Fidler, V. *et al.* (1995). Perinatal exposure to polychlorinated biphenyls and dioxins and its effects on neonatal neurological development. *Early Hum Dev. 41*:111-127.

Hunter, D., Hankinson, S., Laden, F., *et al.* (1997). Plasma organochlorine levels and the risk of breast cancer. *NEJM 337(18)*:1253-1258.

International Agency for Research on Cancer (IARC) (2001). *Overall Evaluations of Carcinogenicity to Humans. Group 1: Carcinogenic to humans.* Available at http://193.51.164.11/monoeval/crthgr01.html at May 7, 2001.

IRIS (2001). Integrated Risk Information Service. Records for *Polychlorinated Biphenyls*, *Aroclor 1016*, *Aroclor 1248*, and *Aroclor 1254*. U.S. Environmental Protection Agency.

Available at http://www.epa.gov/iris/subst/0294.htm (PCBs), 0462.htm (Aroclor 1016), 0649.htm (Aroclor 1248), and 0389.htm (Aroclor 1254).

Israeli, M., and Nelson, C.B. (1992). Distribution and expected time of residence for U.S. households. *Risk Analysis 12*:65–72.

Jacobson, J.L. and Jacobson, S.W. (1996). Intellectual impairment in children exposed to polychlorinated biphenyls in utero. *NEJM 335*:783-789.

Jacobson, J.L., Jacobson, S.W., and Humphrey, H.E. (1990a). Effects of in utero exposure to polychlorinated biphenyls and related contaminants on cognitive functioning in young children. *J. Pediatr. 116(1)*:38-45.

Jacobson, J.L. and Jacobson, S.W. (1990b). Effects of exposure to PCBs and related compounds on growth and activity in children. *Neurotoxicol. Teratol. 12*:319-326.

Karasek, G.L.B (1998). *Hunting Results, Michigan. Small Game Seasons, 1992-1996.* Michigan Department of Natural Resources, Wildlife Division Report No. 3285. February, 1998. Obtainable from (at May 11, 2001) http://www.dnr.state.mi.us/pdfs/hunting/smallgame_9296results.pdf.

Keenan, R.E., and Stickney, J.A. (1996). Letter dated June 24 to J. Cogliano, U.S. EPA, Washington.

Kimbrough, R., Doemland, M., and LeVois, M. (1999). Mortality in male and female capacitor workers exposed to polychlorinated biphenyls *J. Occup. Environ. Med. 41(3)*:161-71.

Kimbrough, R.D., Linder, R.E., and Gaines, T.B. (1972). Morphological changes in livers of rats fed polychlorinated biphenyls: light microscopy and ultrastructure. *Arch. Environ. Health 25*:354–364.

Kimbrough, R.D., Squire, R.A., Linder, R.E., Strandberg, J.D., Montali, R.J., and Burse, V.W. (1975). Induction of liver tumors in Sherman strain female rats by polychlorinated biphenyls Aroclor 1260. *J. Natl. Cancer Inst. 55*:1453–1459.

Kissel, J.C., Richter, K.Y., and Fenske, R.A. (1996). Field measurements of dermal soil loading attributable to various activities: Implications for exposure assessment. *Risk Analysis 16(1)*:115–125.

Koopman-Esseboom, C., Weisglas-Kuperus, N., de Ridder, R.A. *et al.* (1996). Effects of polychlorinated biphenyl/dioxin exposure and feeding type on infants' mental and psychomotor development. *Pediatrics. 97*:700-706.

Knuth, D.E. (1998). *The art of computer programming, Volume 2: Seminumerical Algorithms. Third Edition.* Addison Wesley Longman (ISBN 0-201-89684-2).

Krieger, N., Wolff, M., Hiatt, R., *et al.* (1994). Breast cancer and serum organochlorines: a prospective study among white, black, and Asian women. *JNCI 86(8)*:589-599.

Land, C.E. (1971). Confidence Intervals for Linear Functions of the Normal Mean and Variance. *Ann. Math. Stat. 43*:1187–1205.

Land, C.E. (1973). Standard Confidence Limits for Linear Functions of the Normal Mean and Variance. *J. Amer. Statist. Assoc. 68*: 960–963.

Land, C.E. (1974). Confidence interval estimation for means after data transformations to normality. *J. Amer. Statist. Assoc. 69*:795–802.

Land, C.E. (1975). Tables of Confidence Limits for Linear Functions of the Normal Mean and Variance. *Selected Tables in Mathematical Statistics, Volume III*, 385–419.

Land, C.E. (1988). Hypothesis Tests and Interval Estimates," in *Lognormal Distributions, Theory and Applications*, E.L. Crow and K. Shimizu, eds. Marcel Dekker.

Lanting, C.U., Patandin, S., Fidler, V., Weisglas-Kuperus, N., Sauer, P.J., Boersma, E.R., and Touwen, B.C. (1998). Neurological condition in 42-month-old children in relation to pre- and postnatal exposure to polychlorinated biphenyls and dioxins. *Early Hum. Devel. 50*:283-292.

Lee, C-K., and E-H. Lee (1985). Heat stability of organochlorine pesticide residues in loach. *Bull. Nat. Fish. Univ. Pusan 25*:85–91.

Lentner, C., Ed. (1984) *Geigy Scientific Tables Volume 3: Physical Chemistry, Composition of Blood, Hematology, Somatometric Data*, Ciba-Geigy Ltd., Basle, Switzerland. (ISBN 0-914168-52-5).

Longnecker, M.P., Gladen, B.C., Patterson, Jr., D.G, *et al.* (2000). Polychlorinated biphenyl (PCB) exposure in relation to thyroid hormone levels in neonates. *Epidemiology 11*:249-254.

Loomis, D., Browning, S., Schenck, A., *et al.* (1997). Cancer mortality among electric utility workers exposed to polychlorinated biphenyls. *Occup. Environ. Med. 54*:720-728.

LTI (2000). Limno-Tech, Inc. Remedial Investigation Report and database of water measurements

Lyon, B.F., and C.E. Land (1999). *Computation of Confidence Limits for Linear Functions of the Normal Mean and Variance.* Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6285. ORNL/TM-1999/245. (Available at http://falcon.sis.utk.edu/ConfLimit/ at 9/6/2000).

Mayes, B.A., E.E. McConnell, B.H. Neal, M.J. Brunner, S.B. Hamilton, T.M. Sullivan, A.C. Peters, M.J. Ryan, J.D. Toft, A.W. Singer, J.F. Brown, Jr., R.G. Menton, and J.A. Moore (1998). Comparative carcinogenicity in Sprague-Dawley rates of the polychlorinated biphenyl mixtures Aroclors 1016, 1242, 1254, and 1260. *Toxicol. Sci. 41*:62–76.

MDPH (1987). *Analytical Report of Soil Samples from Lakewood Area, Kalamazoo.* September 23, 1987.

Mendola, P., Buck, G.M., Vena, J.E., *et al.* (1995). Consumption of PCB-contaminated sport fish and risk of spontaneous fetal death. *Environ. Health Persp. 103*:498-502.

Mendola, P., Buck, G.M., Sever, L., *et al.* (1997). Consumption of PCB-contaminated sport fish and shortened menstrual cycle length. *Am. J. Epidemiol. 146*:955-960.

MiCPHA (2000a). Klaviter, E., Humphrey, H., Bloomer, A.W., and Welch, R. *Kalamazoo River Angler Survey and Biological Testing Study, Final Report.* Submitted by Environmental Epidemiology Division, Community Public Health Agency, State of Michigan. Printed by ATSDR. May 2000. A scanned version of this document is included in the supplementary electronic material as Kalamazoo_River_Angler_Survey.pdf (see Appendix B.3).

MiCPHA (2000b). *Kalamazoo River Angler Survey and Biological Testing Study, Database for Phase I.* Provided by the Environmental Epidemiology Division, Community Public Health Agency, State of Michigan, in response to a Freedom of Information Request. Cover letter dated August 31, 2000, from Dr. Robert L. Wahl (Environmental Epidemiology Division) to Dr. Edmund Crouch (Cambridge Environmental Inc.). The original data supplied in response to this request is included in the supplementary electronic information as the file ANGLER2.SD2 (see Appendix B.12).

MiCPHA (2000c). *Kalamazoo River Angler Survey and Biological Testing Study, Database for Phase II.* .Provided by the Community Public Health Agency, State of Michigan, in response to a Freedom of Information Request. Cover letter dated December 23, 1998, from Dr. David R. Wade (Division of Environmental Epidemiology) to Ms. Dawn E. Penniman (Blasland, Bouck & Lee, Inc.). The original data supplied in response to this request is included in the supplementary electronic material in the file Phase_2.zip (see Appendix B.18).

MiDEQ (2000). *Final Human Health Risk Assessment Allied Paper, Inc./Portage Creek/Kalamazoo River Superfund Site.* Michigan Department of Environmental Quality Environmental Response Division. August 18, 2000.

MiIC (2000). Michigan Information Center, Michigan State Budget Office website, www.state.mi.us/dmb/mic/census/demo.htm

Miller, D.T., Condon, S.K., Kutzner, S., *et al.* (1991). Human exposure to polychlorinated biphenyls in Greater New Bedford, Massachusetts: a prevalence study. *Arch. Environ. Contam. Toxicol. 20*:410-416.

Moore, J.A., Hardisty, J.F., Banas, D.A., and Smith, M.A. (1994). A comparison of liver tumor diagnoses from seven PCB studies in rats. *Regul. Toxicol. Pharmacol. 20*:362–370.

Moysich, K., Ambrosone, C., Vena, J., *et al.* (1998). Environmental organochlorine exposure and postmenopausal breast cancer risk. *Cancer Epidemiol. Biomarkers Prev. 7*:181-188.

NCI (1978). National Cancer Institute. *Bioassay of Aroclor 1254 for possible carcinogenicity.* Carcinogenesis Tech. Rep. Ser. No. 38.

Norback, D.H., and Weltman, R.H. (1985). Polychlorinated biphenyl induction of hepatocellular carcinoma in the Sprague-Dawley rat. *Environ. Health Perspect. 60*:97–105.

Patandin, S., Koopman-Esseboom, C., de Ridder, M.A., Weisglas-Kuperus, N., and Sauer, P.J. (1998) Effects of environmental exposure to polychlorinated biphenyls and dioxins on birth size and growth in Dutch children. *Pediatr Res. 44*(4):538–45.

Patandin, S., Lanting, C.I., Mulder, P.G., Boersma, E.R., Sauer, P.J., and Weisglas-Kuperus, N. (1999) Effects of environmental exposure to polychlorinated biphenyls and dioxins on cognitive abilities in Dutch children at 42 months of age. *J. Pediatr. 134*(1):33–41.

Pieters, M.N., Kramer, H.J., and Slob, W. (1998). Evaluation of the uncertainty factor for subchronic-to-chronic extrapolation: statistical analysis of toxicity data. *Regul. Toxicol. Pharmacol. 27*:108–111.

Pluim, H.J., van der Goot, M., Olie, K, *et al.* (1996). Missing effects of background dioxin exposure on development of breast-fed infants during the first half year of life. *Chemosphere. 33*:1307-1315.

Puffer, H.W., and Gossett, R.W. (1983). PCB, DDT, and Benzo(a)pyrene in raw and pan-fried White Croaker (*Genyonemus lineatus*). *Bull. Environm. Contam. Toxicol. 30*:65–73.

Reinert, R.E., Stewart, D., and Seagran, H.L. (1972). Effects of dressing and cooking on DDT concentrations in certain fish from Lake Michigan. *J. Fish. Res. Board Can. 29*:525–529.

Ries, L.A.G., Eisner, M.P., Kosary, C.L., Hankey, B.F., Miller, B.A., Clegg, L., and Edwards, B.K. (eds). SEER Cancer Statistics Review, 1973-1998, National Cancer Institute. Bethesda, MD, 2001. Available at http://seer.cancer.gov/Publications/CSR1973_1998/.

Rogan, W.J. and Gladen, B.C. (1991). PCBs, DDE, and child development at 18 and 24 months. *Ann. Epidemiol. 1*:407-413.

Rogan, W.J., Gladen, B.C., McKinney, J.D., *et al.* (1986). Neonatal effects of transplacental exposure to PCBs and DDE. *J. Pediatr. 109*:335-341.

Royston, J.P. (1982). Algorithm AS 181: The W test for normality. *Appl. Statist. 31*:176–180.

Royston, P. (1993). A toolkit for testing for non-normality in complete and censored samples. *Statistician 42*:37–43.

Royston, P. (1995). Remark AS R94. A remark on algorithm AS 181: The W-test for normality. *Appl. Statist. 44*:547–551.

Schaeffer, E., Greim, H., and Goessner, W. (1984). Pathology of chronic polychlorinated biphenyl (PCB) feeding in rats. *Toxicol. Appl. Pharmacol. 75*:278–288.

Schantz, S., Levin, E., Bowman, R., *et al.* (1989). Effects of perinatal PCB exposure on discrimination-reversal learning in monkeys. *Neurotoxicol. Teratol. 11*:243-250.

Schantz, S., Levin, E., and Bowman, R. (1991). Long-term neurobehavioral effects of perinatal polychlorinated biphenyl (PCB) exposure in monkeys. *Environ. Toxicol. Chem. 10*:747-756.

Schecter, A., Dellarco, M., Päpke, O., and Olson, J., (1998). A comparison of dioxins, dibenzofurans and coplanar PCBs in uncooked and broiled ground beef, catfish, and bacon. *Chemosphere 37*:1723–1730.

Singh, A.K., Singh, A., and Engelhardt, M. (1997). *The lognormal distribution in environmental applications*. EPA/600/R-97/006. December 1997. Available at http://www.epa.gov/esd/pdf/lognor.pdf.

Sinks, T., Steele, G., Smith, A.B., *et al.* (1992). Mortality among workers exposed to polychlorinated biphenyls. *Am. J. Epidemiol. 136(4)*:389-398.

Skea, J.C., Simonin, H.A., Harris, E.J., Jackling, S., Spagnoli, J.J., Symula, J., and Colquhoun, J.R. (1979). Reducing levels of Mirex, Aroclor 1254, and DDE by trimming and cooking Lake Ontario Brown Trout (Salmo Trutta Linnaeus) and Smallmouth Bass (Micropterus Dolomieui lacepede). *J. Great Lakes Res., Internat. Assoc. Great Lakes Res. 5*:153–159.

Smith, W.E., Funk, K., and Zabik, M.E. (1973). Effects of cooking on concentrations of PCB and DDT compounds in Chinook (Oncorhynchus tshawytscha) and coho (O. kisutch) salmon from Lake Michigan. *J. Fish. Res. Board Can. 30*:702–706.

Stanek, E.J. 3rd, Calabrese, E.J., Barnes, R., and Pekow, P. (1997). Soil ingestion in adults--results of a second pilot study. *Ecotoxicol. Environ. Saf. 36(3)*:249–57.

Stehr-Green, P., Burse, V.W., and Welty, E. (1988). Human exposoure to polychlorinated biphenyls at toxic waste sites: investigations in the United States. *Arch. Environ. Health 43(6)*:420-424.

Stewart, P., Reihman, J., Lonky, E., Darvill, T., and Pagano, J. (2000). Prenatal PCB exposure and neonatal behavioral assessment scale (NBAS) performance. *Neurotoxicol. Teratol. 22(1)*:21-9.

Travis, C.C., Richter, S.A., Crouch, E.A.C., Wilson, R., and Klema, E.D. (1987). Cancer risk managementa review of 132 federal regulatory dicisions. *Environ. Sci. Technol. 21*:415–420.

Trotter, W.J., Corneliussen, P.E., Laski, R.R, and Vannelli, J.J. (1989). Levels of polychlorinated biphenyls and pesticides in bluefish before and after cooking. *J. Assoc. Off. Anal. Chem. 72*:501–503.

Tryphonas, H. (1998). The impact of PCBs and Dioxins on children's health: immunological considerations. *Can. J. Public Health 89(Suppl 1)*:S49-52.

Tryphonas, H., Hayward, S., O'Grady, L., Loo, J.C., Arnold, D.L., Bryce, F., and Zawidzka, Z.Z. (1989) Immunotoxicity studies of PCB (Aroclor 1254) in the adult rhesus (Macaca mulatta) monkey — preliminary report. *Int J Immunopharmacol. 11(2)*:199-206.

Tryphonas, H., Luster, M.I., White, K.L. Jr., Naylor, P.H., Erdos, M.R., Burleson, G.R., Germolec, D., Hodgen, M., Hayward, S., and Arnold, D.L. (1991a). Effects of PCB (Aroclor 1254) on non-specific immune parameters in rhesus (Macaca mulatta) monkeys. *Int J. Immunopharmacol. 13(6)*:639–648.

Tryphonas, H., Luster, M.I., Schiffman, G., Dawson, L.L., Hodgen, M., Germolec, D., Hayward, S., Bryce, F., Loo, J.C., Mandy, F., and Arnold, D.L. (1991b). Effect of chronic

exposure of PCB (Aroclor 1254) on specific and nonspecific immune parameters in the rhesus (Macaca mulatta) monkey. *Fundam. Appl. Toxicol. 16(4)*:773–786.

U.S. EPA (1992a). *Dermal exposure assessment: principles and applications. Interim Report.* EPA/600/8-91/011B. January 1992.

U.S. EPA (1992b). Final Guidelines for Exposure Assessment. EPA/600/Z-92/001. 57FR22888–22938, May 29. 2

U.S. EPA (1996). *PCBs: Cancer Dose-Response Assessment and Application to Environmental Mixtures.* National Center for Environmental Assessment, Office of Research and Development, U.S. Environmental Protection Agency, Washington, DC. EPA/600/P–96/001F, September 1996.

U.S. EPA (1997). *Exposure Factors Handbook.* EPA/600/P-95/002Fa, August 1997.

U.S. EPA (1999). Risk Assessment Guidance for Superfund: Volume 3 - Part A, Process for Conducting Probabilistic Risk Assessment. DRAFT. Available at http://www.epa.gov/superfund/pubs.htm#r

Vulsman, T. (2000). Impact of exposure to maternal PCBs and dioxins on the neonate's thyroid hormone status. *Epidemiology 11(3)*:239-241.

Wester, R.C., Maibach, H.I., Sedik, L., Melendres, J., and Wade, M. (1993). Percutaneous absorption of PCBs from soil: in vivo Rhesus monkey, in vitro human skin, and binding to powdered human stratum corneum. *J. Toxicol. Environm. Health. 39*:375–382.

Wilson, N.D., Shear, N.M., Paustenbach, D.J., and Price, P.S. (1998). The effect of cooking practices on the concentration of DDT and PCB compounds in the edible tissue of fish. *J. Exp. Anal. Environ. Epidem. 8*:423–440.

Wolff, M., Toniolo, P., Lee, E., *et al.* (1993). Blood levels of organochlorine residues and risk of breast cancer. *JNCI 85(8)*:648-652.

Yassi, A., Tate, R., and Fish, D. (1994). Cancer mortality in workers employed at a transformer manufacturing plant. *Am. J. Ind. Med. 25*:425-437.

Zabik, M.E., Hoojjat, P., and Weaver, C.M. (1979). Polychlorinated biphenyls, dieldrin, and DDT in Lake Trout cooked by broiling, roasting or microwave. *Bull. Environm. Contam. Toxicol. 21*:136–143.

Zabik, M.E., Merrill, C., and Zabik, M.J. (1982). PCBs and other xenobiotics in raw and cooked carp. *Bull. Environm. Contam. Toxicol. 28*:710–715.

Zabik, M.E., Zabik, M.J., Booren, A.M., Daubenmire, S., Pascall, M.A., Welch, R., and Humphrey, H. (1995a). Pesticides and total polychlorinated biphenyls residues in raw and cooked walley and white bass harvested from the Great Lakes. *Bull. Environm. Contam. Toxicol. 54*:396–402.

Zabik, M.E., Zabik, M.J., Booren, A.M., Nettles, M., Song, J-H, Welch, R., and Humphrey, H. (1995b). Pesticides and total polychlorinated biphenyls in Chinook Salmon and Carp Harvested from the Great Lakes: Effects of skin-on and skin-off processing and selected cooking methods. *J. Agric. Food Chem. 43*:993–1001.

Zabik, M.E., Booren, A., Zabik, M.J., Welch, R., and Humphrey, H. (1996). Pesticide residues, PCBs, and PAHs in baked, charbroiled, salt boiled, and smoked Great Lakes lake trout. *Food Chem. 55*:231–239.

Zabik, M.E., and Zabik, M.J. (1999). Polychlorinated biphenyls, polybrominated biphenyls, and dioxin reduction during processing/cooking food. *Adv. Exp. Med. Biol. 459*:213–231.

Appendix A Absorption of PCBs in the gut

A.1 Analysis for the gut

In the HHRA (MiDEQ, 2000), it was assumed that the fraction of PCBs present in soil that would be absorbed across the gut was unity. This absorption fraction was used as a relative absorption, to be compared with the absorption from the diet occurring in the bioassay that had been used to estimate the carcinogenic potency of Aroclor 1260. No reference is given for this absorption fraction, although U.S. EPA has previously used a value of 0.75 for the relative absorption fraction (again, without reference to any experiment) in developing its advisory levels for PCB clean-up (U.S. EPA, 1986).

Fries *et al.* (1989) examined the absorption in adult male Sprague-Dawley rats of ^{14}C-labelled 2,2',5-trichlorobiphenyl (Tr), 2,2',5,5'-tetrachlorobiphenyl (Te), and 2,2',4,5,5'-pentachlorobiphenyl (Pe) spiked into soil and allowed to age eight years at –5°C. Four experimental procedures were used. In the first, spiked soil was added to the normal diet ("soil diet"); in the second, spiked soil was fed by gavage in a water suspension ("gavage soil"); in the third, the ^{14}C-labelled compound was spiked into the normal diet ("normal diet"); and in the fourth, the ^{14}C-labelled compound was fed by gavage in a corn oil vehicle ("corn oil gavage"). For Tr, only the first two procedures were used, due to lack of ^{14}C-labelled compound.

The experiment that had been used to estimate the carcinogenic potency of PCBs involved the feeding to Sprague-Dawley rats of Aroclor 1260 spiked into the diet. The first ("soil diet") and third ("normal diet") experimental procedures of Fries *et al.* (1989) thus allow a direct evaluation of relative absorption for the situation most relevant on this Site. The computed relative absorption for the two compounds was:

2,2',5,5'-tetrachlorobiphenyl	0.81 to 0.87, and
2,2',4,5,5'-pentachlorobiphenyl	0.68 to 0.78

The ranges cover the uncertainty due to the inability of Fries *et al.* (1989) to distinguish what fraction of the un-extractable material could be the compound of interest in both the original material and in the rats' feces.

As expected, there appears to be lower absorption for the more highly chlorinated PCB, and weighting these absorption fractions by the fractions of these or more extreme chlorination levels (26% Cl_4 or lower, 74% Cl_5 or higher, for the average Aroclor composition of the impoundment surface soil samples) gives a range of relative absorptions of 0.72 to 0.80. A rounded mean estimate of 0.76 was therefore used in this report for the required relative absorption fraction.

A.2 References for this appendix

Fries, G.F., G.S. Marrow, and C.J. Somich (1989). Oral bioavailability of aged polychlorinated biphenyl residues contained in soil. *Bull. Environ. Contam. Toxicol.* 43:683–690.

MiDEQ, 2000. Final Human Health Risk Assessment Allied Paper, Inc./Portage Creek/Kalamazoo River Superfund Site. Michigan Department of Environmental Quality Environmental Response Division. August 18, 2000.

Appendix B Spreadsheet calculation details

B.1 Introduction and supplemental spreadsheets

The calculations for this risk assessment were performed in spreadsheets and in a special computer program (see Appendix C). The spreadsheets and program were designed to run in Windows® 98 (Microsoft Corporation) in Intel iAPX® processors with co-processors (Pentium or later); although all should run in versions of Windows® from 95 onwards, they have not been tested for any other configuration. All the spreadsheets are part of the electronic supplemental information material accompanying this risk assessment. All are Quattro® Pro 8 (Corel Corporation) spreadsheets. As explained in Section B.2, many of these cannot be translated to other spreadsheet formats without substantial effort to reproduce special add-in functions (and some of the spreadsheet techniques used may not translate well either). The spreadsheets contain cross-references to one another, so should all be available in the same sub-directory.[1]

Table B.1 lists the spreadsheets and cross-references where in the Risk Assessment they are used and where they are discussed in this appendix. The following sections give some guidance on their contents and construction. Cells and cell ranges (blocks) in the spreadsheets have been extensively named, so that the discussion is often largely in terms of the block names defined in the spreadsheets (rather than cell references). A list of the names defined in a spreadsheet is available by clicking on the "Navigate" button (a stylized pointing hand) on the input line; clicking on a name on that list will select the named cell(s).

Very extensive use has been made of the optimizer available in Quattro Pro 8 to obtain maximum likelihood estimates, likelihood-based confidence bounds on those estimates, and estimates of information matrices. For the convenience of those familiar with these techniques, but unfamiliar with the use of spreadsheets to implement them, some examples of the optimizer set-ups have been stored in named blocks within the spreadsheets.

[1] On opening linked spreadsheets, one will be prompted whether to open the supporting spreadsheets also. To prevent subsequent errors, it is advisable to initially open all supporting spreadsheets. They can then be closed, and links updated manually only when necessary. Experience indicates that Quattro Pro 8 may experience errors if the supporting spreadsheets are not opened at least once before attempting to manually update links.

The spreadsheets contain calculations in various units. The units for a given number are generally given in the cell adjacent to (or above, for tables) the number. Some of these units are expressed using the Greek μ symbol (for the micro = 10^{-6} prefix) using the Symbol font, while in other cases the symbol u is used for this purpose. If the Symbol font formatting is lost for some reason, the resulting unit will probably incorrectly display as m (milli = 10^{-3}).

Table B.1 Spreadsheets and other files used in this risk assessment and provided in the supplemental information.

Spreadsheet or other file Name	**Used in Sections**	**Discussed in Section**
EACC_functions.zip	Used in spreadsheets	B.2
Kalamazoo_River_Angler_Survey.pdf FOIA_requests_responses.pdf	For reference	B.3
PCB_Cancer_dose_response.wb3	4.2	B.5
Land_Table.wb3 and Land_Lyon.wb3	Referenced from other spreadsheets	B.6
Impoundment_data.wb3	5.1	B.7
Fish_data_HHRA.wb3	6.2.1	B.8
Bass_Carp_time.wb3	6.2.4	B.9
Fish_data.wb3	6.2.2, 6.2.5.1, 6.2.5.2, 6.2.5.3, and 6.2.7	B.10
Other_exposure.wb3	5.2, 5.3, 6.2.4, 7.1, 7.2	B.11
Phase_1.wb3 (and ANGLER2.SD2)	6.2.6, 6.3.1	B.12
PCB_congener_data.wb3	4.3.1, 6.3.2, 7.2	B.13
Age_structure.wb3	6.4	B.14
Meals.wb3	6.3.1, 6.5.1, 6.5.3	B.15
Cooking_effect.wb3	6.6.1	B.16
Atkin_survey.wb3 and Fish_Codebook.doc	6.5.3, 6.6.2	B.17
Phase_2.wb3 (and Phase_2.zip)	6.6.2, 6.11.2	B.18
Surface_water.wb3	7.2	B.19
Examples.wb3	6.10.4	B.20

Dose_life_results.wb3	6.10.1, 6.10.2, 6.10.3, 6.10.5, 6.11.1, 6.11.1	B.21
Dose_while_results.wb3	6.10.1, 6.10.2, 6.10.3, 6.11.1, 6.11.1	B.22
Risk_results.wb3	6.10.4, 6.10.5	B.23
HI_results.wb3	6.10.4	B.24

B.2 EACC_functions.zip — special spreadsheet add-in functions (@EACC library)

Many of the spreadsheets contain special add-in functions to perform complex calculations. Quattro® Pro 8 allows the user to load custom-designed @function libraries (compiled into dynamic link libraries, DLLs) either automatically, or by using a macro command ({DLL.Load}), or by inserting such a command as the value of the InitMacro key in the Registry. Automatic loading works correctly, and is all that has been tested. Automatic loading requires that the DLL be present in the directory in which QPW.EXE resides, and also requires any required compiler run-time library files to be in the C:\WINDOWS\SYSTEM directory (if Windows is installed elsewhere, substitute the appropriate location, although no alternative has been tested). Use of the special add-in functions then simply requires entry of the fully qualified name for the function into a formula, as usual (*e.g.* @EACC.pNorm(b25)), and the DLL is automatically loaded (if it was not already loaded). Subsequently (during that session), you can just enter the unqualified function name, and the DLL name qualification will be added automatically. The function names given in the documentation (see below) correspond to those that are shown in Quattro Pro 8 after you have entered them. Function name entry in Quattro Pro 8 is not case sensitive, so you can type the functions in any way you please. After a library has loaded, you can see (somewhat cryptic) prompts for the arguments at the bottom right, as usual for all Quattro functions.

The special function add-in library used in these spreadsheets is named @EACC.DLL. This DLL is included in the file EACC_functions.zip, which contains the following files:

eacc.dll	The DLL with the @EACC special functions.
eacc.cpp	The source code for eacc.dll.
eacc.def	The .def file for eacc.dll.
Myfuns.pdf	Documentation for the @EACC special functions.
bds52f.dll	Borland re-distributable run-time library file (see below).
cw3230.dll	Borland re-distributable run-time library file (see below).
owl52f.dll.	Borland re-distributable run-time library file (see below).

The available functions are documented in Myfuns.pdf, and the C++ source code is included in the file eacc.cpp. Construction, compilation and use of such special @function libraries is described in the Software Development Kit (SDK) accompanying the Corel WordPerfect Suite 8 software (it is usually in the \SDK directory on the CD-ROM). According to the SDK documentation, compilation requires the eacc.def file included above (although this apparently adds errors to the link stage of compilation, the resultant DLL works as desired). The source code was compiled using the Borland C++ 5.02 compiler. The Borland C++ 5.02 re-distributable run-time library files are named bds52f.dll, cw3230.dll, and owl52f.dll.

Full use of the spreadsheets requires the following actions before Quattro Pro is started:

- Copy the Borland C++ 5.02 re-distributable run-time library files bds52f.dll, cw3230.dll, and owl52f.dll to the WINDOWS\SYSTEM directory.
- Copy the @function library EACC.DLL to the directory containing QPW.EXE (by default, C:\Corel\Suite8\Programs).

Subsequently, the special @functions should be available as described.

It is expected that the add-in functions would work correctly in Quattro Pro 9 (part of the Corel Office 2000 suite), with appropriate changes in file locations. However, this has not been tested, and use of Quattro Pro 9 is not recommended — duplication of many of the spreadsheet analyses requires extensive use of the optimizer, and the manual optimizer interface in Quattro Pro 9 has known and unfixed bugs.

The use of these special functions makes translating the spreadsheets to other formats (*e.g.* Excel) problematic. None of the special functions will translate (in addition to the problems that arise in translation normally), and they would have to be specifically ported to the new spreadsheet.

The use of special functions are essential to some of the analyses; and such analyses are impractical in spreadsheets without the special function add-ins.

B.3 Kalamazoo_River_Angler_Survey.pdf and FOIA_requests_responses.pdf

The first is a scanned image of the Kalamazoo River Angler Survey (MiCPHA, 2000a), including the questionnaires for the Phase I and Phase II parts of the survey. The best available copy was scanned, but pages 64 and Appendix J, page 5, were missing.

The second contains scanned images of the FOIA requests for the raw survey data, and the responses provided (MiCPHA, 2000b,c).

Both are provided for reference purposes, since some of the information included is not published elsewhere.

B.4 Standard analysis (concentration data)

For many sets of concentration data, a standard analysis is performed that corresponds to a default EPA-style analysis of environmental concentrations. This analysis is implemented through a standard piece of spreadsheet code that can be copied from place to place (the code uses some add-in @functions, see Section B.2) . To operate correctly, sample concentration data have to be present in a single column, sorted in increasing order downwards, with the logarithms of the sample data in the column next to the right. The standard block of code is then copied three columns to the right of the top sample concentration entry (two columns to the right of the logarithm column, so there is a single blank column between the logarithms and the standard code block), and the correct number of samples entered in the appropriate place in the spreadsheet code (most easily by using the @count function). The block of spreadsheet code produces the following set of statistics:

Number of samples (entered manually, or set up with the @count function applied to the samples)
Mean concentration.
Sample standard deviation of concentration.
Mean of natural logarithm of concentration.
Sample standard deviation of natural logarithm of concentration.
Shapiro Wilk statistic testing the sample set for normality.
Shapiro Wilk statistic testing the sample set for lognormality.
t-statistic for these samples.
H-statistic for these samples (obtained as explained in Section B.6).
95th percent upper confidence limit (UCL95) on the mean, assuming a normal distribution of samples, using the t-statistic (normal estimate).
UCL95 on the mean, assuming a lognormal distribution of samples, using the H-statistic (lognormal estimate).
Maximum of the sample values.
Selection of a distribution type — "Lognormal" if the Shapiro Wilk statistic for lognormality is less than 0.05; otherwise "Normal" if the Shapiro Wilk statistic for normality is less than 0.05; otherwise "Neither."
Choice of the estimate for the UCL95 on the mean; lognormal estimate if "Lognormal" is selected, otherwise the normal estimate.
Selected upper bound estimate for mean concentration: the maximum measured value if the chosen UCL95 estimate exceeds that maximum, otherwise the chosen UCL95 estimate.

The object of the code is to produce an upper bound estimate of the mean for the sample concentration dataset. This upper bound is produced using the algorithm:

a) Compute the Shapiro-Wilk statistic ((Royston, 1982, 1993, 1995) for the logarithms of the concentration data.

b) If the Shapiro-Wilk statistic computed in a) exceeds 0.05 (data are consistent with a lognormal distribution), estimate the UCL95 on the mean using Land's procedure (Land, 1971, 1973, 1974, 1975, 1988; Lyon & Land, 1999). Otherwise estimate the UCL95 on the mean using the t-statistic (as though the data are normally distributed).
c) Select the lower of the UCL95 estimate and the maximum of the measurements.

B.5 PCB_cancer_dose_response.wb3

The main calculations are performed on the sheet labeled Calcs, which operates on the experimental data. The optimization procedure has been stored in Optim_model — the "Load model" option can be used to recover it. Other non-default options were to set the precision to 1E-09 and the tolerance to 1E-08. The optimization maximizes the loglikelihood with respect to the parameters of the dose-response curves and the three ratios (R_16_54, R_42_54, and R_60_54) of potency estimates, while constraining some ED_{10} values (expressed as functions of the dose-response parameters and located in ED10_constraints) to correspond to a 10% increment in probability of cancer. Scaled versions of the parameters (in scaled_params and scaled_ratios, respectively) are optimized — such scaling improves the convergence, stability, and accuracy of the optimization with the numerical optimizers used in such spreadsheets. The eight unconstrained ED_{10} estimates correspond to the eight independent estimates available, while the seven constrained values correspond to the requirement that the relative potencies of the different Aroclors be fixed. Despite the un-symmetric implementation, the solution is entirely symmetric — it makes no difference to the results which of the fifteen experiments are chosen to be the independent ones.

There is a switch that allows treating Clophen A-30 as Aroclor 1016 or 1242. The selected values is 1016. Scaled optimum parameters with the other choice have been stored in columns Q (for the dose-response parameters) and AB (for the ratios of potencies). However, the maximum loglikelihood with this choice is very much lower than with the choice of 1016.

The sheet Extra_info shows the calculation of average dose rate for Kimbrough *et al.* (1975), and the comparison between homolog profiles for Clophen A-30 and A-60 and the Aroclors.

The sheet Sensitivity evaluates the relative variation in overall carcinogenic potency arising from a factor two variation in the relative potencies of the other Aroclors relative to Aroclor 1254.

B.6 Land_Table.wb3 and Land_Lyon.wb3

The spreadsheet Land_Table.wb3 is cross-referenced by several others. It contains a table of values of the statistical parameter H_{95} (Gilbert, 1987), obtained from Land (1995), and used in estimating the upper 95th percent confidence limit on the mean of a lognormal distribution. For a sample of size n from a lognormal distribution, with sample mean m and sample standard deviation s, the upper 95th percent confidence limit U is given by:

$$U = \exp\left(m + s^2/2 + sH_{95}/\sqrt{n-1}\right)$$

The values of H used in other spreadsheets are obtained from this table by using 2-dimensional order 3 Lagrange interpolation in ln(s) and ln(n) on ln(H). This was found to be as accurate, or more accurate, than the 4th order interpolation in s and n on H suggested in Land (1995). Note that $n = \nu + 1$, where ν is the number of degrees of freedom used by Land (1995) for tabulation.

The second spreadsheet, Land_Lyon.wb3, has an extensive comparison of the approximate distribution used in the Monte Carlo analysis with the calculated exact distribution of Land (Lyon & Land, 1999). Columns D through X contain the exact values of the H statistic for 19 percentage points (0.25%, 0.50%, 1.00%, 2.50%, 5.00%, 10.00%, 20.00%, 30.00%, 40.00%, 50.00%, 60.00%, 70.00%, 80.00%, 90.00%, 95.00%, 97.50%, 99.00%, 99.50%, 99.75%), for 46 values of standard deviation from 0.1 to 12, and for degrees of freedom 2 through 20 (each value), 22 through 30 (every even value), 35 through 100 (every 5th value), 110 through 200 (every 10th value), 220 through 500 (every 20th value), 550 through 1000 (every 50th value). Columns AA through AS convert the H statistic to the deviation from the mean. Columns AU through BO contain simulation results using the approximation discussed in Appendix C.1.10 for 2 through 20 degrees of freedom (which covers the range required in this document). Columns BQ through CI then show that the simulation results are slightly biased high (approaching unbiased for large degrees of freedom and large standard deviation, as shown by the fraction of simulation results exceeding the exact results approaching 0.5 in those circumstances).

B.7 Impoundment_data.wb3

Sheet labeled All surface.
This contains the former impoundment surface sample data, with self-evident headers, in rows 4 to 203. In rows 209 to 407, any J, B, C, and CJN qualifiers are removed — all such qualifiers are ignored. Rows 413 to 611 extract numerical values from the string values previously processed, entering the negative of the detection limit for non-detects (for subsequent processing). Rows 618 to 816 then combine duplicate entries by averaging them if both are detects, or taking the largest entry if either or both are non-detects. The effect is:

Two detects	Take the average
One detect, one non-detect	Take the detected value
Two non-detects	Set the effective detection limit as the smaller of the two detection limits.

Columns T and U obtain the possible range of total PCB concentration (lower and upper respectively) for each sample. The lower value is the sum of detects of Aroclors 1016, 1242, 1248, 1254, and 1260. The upper value adds the detection limits for non-detected Aroclors in the same list of Aroclors. Aroclors 1221 and 1232 are treated as not present (zero concentration). The average Aroclor fractions (detects only) for all surface soil samples in the former impoundments is obtained at Soil_average_aroclor.

Sheet labeled Calcs.
The range (lower to upper) of possible total PCB concentrations has been copied for all surface samples from Sheet "All surface", grouped by former impoundment, grouped by surface elevation within impoundment as described in Section 5.1. The measurements within each former impoundment have been placed in order of average sample concentration (average of lower and upper).

For the data on each former impoundment, the standard set of statistics (see Section B.4) for the sample concentrations (average of lower and upper estimates) is evaluated in columns AI through AO; some are summarized in the table Standard_Table. An example block of standard statistics is AI28 through AI35.

For the former Plainwell impoundment, where the average sample concentrations data are consistent with coming from a lognormal distribution, two further analyses (likelihood-based, and a jackknife approach, as discussed in Section 5.1) are implemented in columns W through AC.

For all three former impoundments, the likelihood-based approaches using the possible range of concentrations for each sample are implemented in columns J through O (two-component lognormal models for Otsego and Trowbridge, one-component for Plainwell) and P through U (three-component lognormal models for Otsego and Trowbridge, two-component for Plainwell). Parameters in these models are labeled with names that indicate their meaning (m or mu for mean; sigma or s for standard deviation; f or g for the fractional component), and which former impoundment they refer to (_o for Otsego, _t for Trowbridge, _p for Plainwell). Sample optimizer set-ups are in the spreadsheet sheet labeled Optim.

For each of the former impoundments, the maximum likelihood parameters have been copied for quick reference into the column labeled "MLE," and the parameters corresponding to the UCL95 on the mean estimate have been copied to the column labeled "UCL95." A third column, labeled "Graph" contains the set of parameter used to plot the graphs present in the "Objects" sheet of

the spreadsheet (set up in the sheet labeled Graphs). The parameter estimates in this last column have been labeled using names that have an additional _g appended.

Sheet labeled Graphs
Contains the information required to produce the graphs that are Figures 5.1, 5.2 and 5.3.

Sheet labeled Optim
Contains example optimizer setups for fitting the impoundment data distributions and finding upper confidence limits on the mean concentration.

B.8 Fish_data_HHRA.wb3

An analysis of the HHRA fish dataset. These are the measurements on carp and bass in 1993 and 1997 only, with ABSAs 3, 4, and 5 combined. There are two sheets in the spreadsheet, one for carp and one for bass. Only the "total PCB" data are used; all J qualifiers are stripped, and ½ the detection limit is used for reported non-detects. For each ABSA, the standard set of statistics (see Section B.4) is computed, and they are summarized in Carp_Table and Bass_Table, which correspond to Tables 6.1 and 6.2.

B.9 Bass_Carp_time.wb3

The analyses of time trends in bass and carp discussed in Section 6.2.4. The sheets labeled Bass and Carp contain total PCB fish data, obtained from spreadsheet Fish_data.wb3 (Section B.10), including lower and upper limits on possible PCB concentration when account is taken of detection limits. Columns V and W contain the model predictions for mean and standard deviations (the model expressions may be constructed automatically by copying the string functions saved at V18 and W18), and column Y the loglikelihood contribution for each fish. The model parameters are all named on the Analysis sheet. Copies of the MLE for all parameters, for equal time trend coefficients in bass and carp, and for equal time trends and zero coefficient for the weight, are shown on the Analysis sheet beneath the working parameters. The likelihood analysis results giving the p-values cited in Section 6.2.4 for these cases are shown within the parameter blocks at C50..D53 and C70..D73 respectively — the method of computation is shown at D20..E23.

The uncertainty distribution for the common time trend is shown in columns P through S of Analysis, a normal plot of it in chart Beta_all, and a normal fit to it using regression analysis in columns U through X. The optimizer program for computing the lower end of this uncertainty distribution (using the profile likelihood approach) is stored in the Optim sheet — the percentile is set in H23 before optimization to find the given percentile. For the upper end of the uncertainty distribution (above the 50th percentile), the optimizer must be set to minimize rather than maximize. The distribution in column Q gives the same sign to the time trend as used in

Section 6.2.4 (positive means decreasing concentration) — the opposite sign is used for the spreadsheet parameter values (negative means decreasing concentration).

B.10 Fish_data.wb3

Lists the data for carp, bass, other fish, and turtles in its original format, strips J and N qualifiers, evaluates detection limits for non-detects, and computes the range of total PCB values (minimum to maximum, with non-detect equal to zero or the detection limit, respectively) for each sample. Performs the standard analyses (see Section B.4) on the average of minimum and maximum PCB concentrations (*i.e.* treating non-detect as ½ the detection limit). At the top of each sheet of the spreadsheet are tables corresponding to Tables 6.3, 6.4, 6.5, 6.6, 6.7, and 6.8 in Sections 6.2.5.1, 6.2.5.2, and 6.2.5.3. The Averages sheet, contains the average Aroclor fractions in each species of fish and turtle, as used in Section 6.2.7. The sheet All_fish contains all the fish data (not turtles) sampled, whether included or not in the risk assessment, examining which Aroclors were ever detected.

B.11 Other_exposure.wb3

The Calculation sheet is a straightforward series of calculations to match the scenarios described.

The Hunter_info sheet reproduces material from Karasek (1998) used in Section 5.2.1, and computes the average times given there.

The Dermal_contact sheet performs the calculations detailed in Section 5.2.1 and resulting in Table 5.2. In addition, the same procedure is followed for farmers, for use in the gardener scenario in Section 5.3.2.

The UCL calculations here are for the distribution of individual events, and so are not relevant for the calculation (which requires the mean over many events).

The Garden sheet derives total PCB concentrations for the garden soil and produce samples, and gives mean produce consumption rates.

The sheet Dam performs the calculations outlined in Section 7.2, and shows the calculation of PCB flow rates in the river (used in Section 6.2.4). Input values are located in A15..E22. Source locations are in B65..B90. Receptor distances are calculated at D65..I90, the wind angle for each source-receptor pair in K65..P90, downwind and crosswind distances in S65..X90 and Z65..AE90, with the resulting values of SigmaZ and SigmaY in AG65..AL90 and AN65..AS90. Then downwind concentrations (unit emission per source and 1m/s windspeed) and uniform-wind-rose averages are in AU65..AZ90 and BB65..BG90, with sums over all sources in AU92..AZ92 and BB92..BG92. Concentrations assuming complete emissions of PCBs are then

in F31..K32, and adjustments to account for Henry's law and windspeed in B95..L128, with the final adjustment for directionality in F36..K36. The maximum concentration at 5 m distance is obtained by maximizing F31 with respect to wind angle in D16, and results have been copied into the table at C38..L55.

Sheet Air_data contains the meteorological data (from the U.S. EPA SCRAM bulletin board) for 1992 from Grand Rapids/Kent County International Airport. The wind speed and direction have been extracted and frequency tables constructed.

B.12 Phase_1.wb3

The Phase I raw data (MiCPHA, 2000b) were obtained from Dr. R.L. Wahl of Michigan Department of Community Health in response to a Freedom of Information Act request. The data were encoded in a SAS® data file "ANGLER2.SD2". The Phase I questionnaire is included as Appendix B of MiCPHA (2000a). Data were extracted to a Quattro Pro spreadsheet using the SAS® Universal ODBC Driver (Version 1.1).

The Raw_data sheet contains the raw data in columns A to DD. Columns DG to DI compute the numbers of fish-eating anglers who were fishing in Kalamazoo and Allegan counties, respectively, as used in Section 6.2.6.

The sheet Years_eating uses column "Caught" of sheet Raw_data. The values in that column have been copied to column B of sheet Years_eating, and sorted. The five zero values are ignored, as was the single value of 80 years. The cumulative distribution, equation 6.8 (see Section 6.3.1) is fitted and Kolmogorov-Smirnoff statistics calculated, for single and double exponential curves.

The sheet s_dates evaluates the days of the week on which anglers were interviewed, and the distinct areas in Allegan county where anglers were interviewed. The calculations detailed in Section 6.7 are performed in AF17..AH36 (probability for an angler to be captured by the survey) and AN17..AT43 (calculation of populations).

Some modifications were made to the survey data where it was clear that a data entry error had been committed, and the correct value was apparent from the context. The following modifications have not discussed elsewhere in the text:

Column	UniqueID	Original value	Corrected value	Reason
County	003642	4	A	Sequence numbering and location columns indicate this is Allegan county
Date	9 entered dates outside the range of the survey corrected by their context. See sheet s_dates of Phase_1.wb3.			

B.13 PCB_congener_data.wb3

The Data sheet contains data on PCB congener composition of Aroclors determined by Frame (1996), as obtained from the U.S. EPA web site (see the Frame *et al.*, 1996 citation). In addition, data on Aroclor 1232 was entered directly from Frame *et al.* (1996) for estimating the Henry's law constant for PCBs dissolved in the Kalamazoo river.
The Dictionary sheet contains the original data dictionary for the database so downloaded.
The Properties sheet contains properties computed from the structure of the PCBs, Henry's law estimates (based on Brunner *et al.*, 1990), and rate constants from Brown (1994). The congener composition of the standard 75% bass, 25% carp mix is computed in column AM, then the accumulated body burdens (equation 6.12) and the time integral of body burdens (equation 6.14) in columns BV..CX and AP..BR respectively. The approximation functions for the NOAEL (equation 4.3) and for h (equation 6.17) are estimated in blocks AP2..AS12 and BV2..BX10 (exact and approximate values are compared in lines 239 to 245). The mix of congeners in water is computed in column AN and used to estimate the Henry's law constant (in H_water). The ~6% unchlorinated biphenyl in Aroclor 1232 (Frame *et al.*, 1996) has been ignored in this approximate calculation.
The sheet Brunner_Exptl reproduces the correlation used by Brunner *et al.* (1990) for Henry's law.
The sheet Graphs sets up graphs used in this report.

B.14 Age_structure.wb3

All calculations are performed in one sheet. The integral of equation 6.18 is approximated by sums using a ½ year step. Column AI gives ½ year values from 0.25 to 84.75, column AJ evaluates $1 - F_2$ at these times, column AK evaluates q at these times, and column AL estimates the integral from zero to a ½ integer year of q by a sum over the ½ yearly values. Columns AM through AQ then estimate the finite s integrals of $q(s-T)$ over the age ranges available from the survey, for ½ yearly steps in T (centered on the 1/4 integer entries in colum AI). Columns AC through AG multiply by the appropriate values of $1 - F_2$, and columns Q through U then sum over the appropriate T periods to match the observations, approximating the T integral. The

entire integral is normalized using normalizer so that the expected and observed number of people in the survey match.

Parameters for the model are listed in K5..L11. For the piecewise linear model used for initial testing, the model used in column AK was replaced with the model example shown in AK11. This has the effect of linearly interpolating into a table of values for q (the parameters of the linear interpolation table are derived from K11..L18, where they are convenient for optimization).

The multinomial loglikelihood parts are computed in columns W through AA. It is possible to evaluate the total loglikelihood using either the individual periods of fish-eating (rows 46 through 101), or with the observed and expected values binned to the period ranges discussed in Section 6.4 (rows 108 through 119). The latter was chosen as the best estimate.

The MLE parameters are given in H5..H9. The second derivative matrix of the likelihood with respect to the parameters is estimated in S4..AG12. A change in the parameters in L5..L9 is shown in Q5..Q9, and twice the corresponding change in likelihood in Q11. Each column in S4..AG12 show the small changes applied first to individual parameters (S..W), then in two parameters at a time. Line 12 computes the second derivatives, which have been copied to the matrix AK5..AO9.

AJ12..AK18 can be used to demonstrate that this matrix is positive definite — AJ12..AJ16 is an arbitrary vector, and AK18 the quadratic form obtained from this vector with the matrix of second derivatives. The optimizer can be used to minimize AK18 over vectors AJ12..AJ16 (with the constraint that they be of unit length, AJ18) and show that it is positive.

AQ5..AU9 is then the inverse of the 2nd derivative matrix, and so is the variance-covariance matrix. The lower-diagonal square-root of the variance-covariance matrix (that is, the lower-diagonal matrix whose product with its own transpose gives the variance-covariance matrix; this is used in the computer program to generate random variates from the uncertainty distribution) is obtained at AW5..BA9.

Finally, AQ11..AU15 obtains the standard deviations (leading diagonal) and Pearson product-moment correlation coefficients (below the leading diagonal) from the variance-covariance matrix.

B.15 Meals.wb3

Sheet Number has copied into it (from Phase_1.wb3) the responses to questions relevant to number of meals per year and length of time the angler has been eating fish (C30..BC969). Columns BE..BN calculate various statistics given in Section 6.5.1. Columns BS..CA obtain estimates of number of meals per year and length of time eating fish. Columns CH through CH have been simultaneously sorted to remove non-responders. They are the raw data used for the model, which is fitted using columns CM..CS. The model is specified by the parameters in CO3..CP6. MLE estimates are in CM3..CM8. The variance-covariance matrix is estimated as described in Section B.14 — small offsets from the MLE may be entered in CQ3..CQ6, and the resulting values and loglikelihood copied (using Convert to Values) to the columns in the block CS3..DB8. The 2nd derivative are estimated in row CS9..DB9, and the rest of the procedure is as described in Section B.14.

The remainder of the sheet sets up the graph to display observed and estimated distributions for meals/year, for each of four quartiles of the length of time eating.

Sheet Size contains the meal-size distribution discussed in Section 6.5.3, with computation of average meal size. It is here re-ordered to put meal sizes in order of their probability, since this is the order used in the computer program. Values from this sheet were copied to the program file.

Sheet Correls computes correlation coefficients between length of time eating fish, number of meals per year, and number of times fished in the last calendar year, using best estimates for these values.

Sheet Fish_frac sets up the data file eat_data.dat of fish-type fractions for each responding angler, separated into approximate quartiles by meals/year.

B.16 Cooking_effect.wb3

Sheet Raw_data contains the measurements presented in each paper, in as much detail as is presented there. Calculations are performed, where necessary and possible, to obtain mean and standard deviation of the measured loss of PCBs for each combination of parameters discussed in the original paper.

In sheet Summary, the data from Raw_data are summarized. Weighted averages are then computed (block AB12..AE27) that are used to form the distributions (graph Distr, set up in AI8..AL32). The means of the distributions, and the calculation of a weighted grand mean, are in columns AN..AS.

B.17 Atkin_survey.wb3

This spreadsheet contains the data from the survey by Dr. Charles Atkin (1994), and an associated file, Fish_Codebook.doc (in Microsoft Word format), contains the codebook. The original data file was a Microsoft Excel spreadsheet; the Raw sheet of Atkin_survey.wb3 is a direct copy from that spreadsheet, with added header rows, but with the 6-digit partial phone numbers completely removed.
The sheet Meal_size computes the average meal sizes for fishers reporting both meal size and number. Also evaluated is are estimates of meal size by fish type.
The Cooking sheet contains frequency tables for whether fish are fried.

B.18 Phase_2.wb3 and Phase_2.zip

The Phase II data (MiCPHA, 2000c) were provided in response to a Freedom of Information Act request by Dr. D.R. Wade of the Michigan Department of Community Health. The data files provided were described as follows:

> *1) Kalex2.xls - this is an Excel file which contains the responses to the one-page exposure history questionnaire. There are 157 observations in this file. Coding is apparent from hard copy of the questionnaire or labeling of the file.*
>
> *2) Angler.dbf: this is foxpro 2.6 data file which contains the laboratory results for 211 individual samples.*
>
> *3) Kalam4.rec: this file is in Epi-Info ver. 5 and contains the questionnaire answers for 156 individuals. Coding for these files can be determined from the hard copy questionnaire included or by browsing the file while in Epi-Info. (Epi-Info is free software available from CDC off their web site www.cdc.gov)*
>
> *Observations from these three files can be linked by id number. No identifying information is available in these files. The disk also contains the following coding documentation in Word Perfect 6:*
>
> *1) Alphcit.kam: City codes for hospital locations.*
> *2) Alphospn.kam: numeric codes for hospitals*
> *3) medcod.kal: medication duration codes*
> *4) medcodal.kam: medication type code*
> *5) repcona.kal: medical conditions codes*

The Phase II questionnaire is included as Appendix C of MiCPHA (2000a), see Appendix B.3. Copies of the original files are included in the supplementary material in the Zip file Phase_2.zip. Data were transferred from Kalex2.xls and Angler.dbf directly to the Quattro Pro spreadsheet

Phase_2.wb3 (those file types may be opened directly in Quattro Pro). The questionnaire data from Kalam4.rec were extracted using Epi-Info 2000 to produce a Microsoft Access Database, which could then be accessed directly from Quattro Pro using the standard ODBC driver. Where necessary, cells with numeric codes that transferred as text were converted to numerical values.

The sheet KALEX2 of Phase_2.wb3 is a copy of the Kalex2.xls file, with the modification noted to one ID number (to match the other files).

The sheet Questionnaire contains the responses to the Phase II questionnaire (recorded in Kalam4.rec), with one ID number modified as noted (to improve the matching between files). Columns GA..GC provide a summary table of the number of fish meals (with fish from the Kalamazoo) eaten in the last 12 months, and whether any fish preference was recorded (Section 6.11.2). Columns GG..GN evaluate serving size (Section 6.5.3)

The sheet IDCompare shows the matching of ID numbers between the three data files.

The sheet Blood evaluates the conversion between PCB body burden and blood concentration (Section 6.11.2)

The sheet Fishdata summarizes concentration measurements for the various fish (from Fish_data.wb3), adjusts them to 1995, and averages them over ABSAs as described in Section 6.11.2.

The sheet Concs contains the information encoded in Angler.dbf, with a modified ID number as noted, and with the concentration entries modified to agree with the latest published values in MiCPHA (2000a) — although one page of the data is missing from MiCPHA (2000a, Appendix J, page 6). The detection levels are from MiCPHA (2000a, Appendix J). 155 of the 211 samples in the file are identifiable with individuals (the others are presumed to be quality control samples). In addition, the models described in Section 6.11.2 are implemented in this sheet. Columns H..W extract and process required data from the other sheets. Columns W..AJ are concerned with the age-only model (equation 6.28), with the parameters in AB5..AD16, and results in AG1..AO22. Columns AL..BI (below row 26) implement the model containing both age and amount of fish eaten (equation 6.30), with parameters in AX5..AZ11 and results in BB1..BG16. Summary statistics are calculated in BA22..BJ25.

B.19 Surface_water.wb3

Sheet LTI_2000 records the measurements of surface water concentrations (LTI, 2000).

Sheet Kzoo_river abstracts all measurements on the Kalamazoo, and estimates the mean and Aroclor composition. Aroclors 1016, 1221 were not detected at all, and treated as not present. Duplicates are combined according to:

Both detects	Take the average.
One detect	Use the detected value.
Both non-detect	Non-detect, with the lowest detection limit of the pair.

Non-detects are treated as ½ the detection limit. Fitting a model where all measurements were samples from a lognormal distribution of total PCB concentrations, with fixed fractions of the ever-detected Aroclors, gives very similar estimates (columns AQ to AV), but this model was not used.

B.20 Examples.wb3

The examples given in Section 6.10.4, Tables 6.20, 6.21, 6.22, and 6.23, together with others, are worked through in this spreadsheet.

Sheet Import is used to import the examples.txt file produced by the Monte Carlo program. Some of the tables are formatted for export to the text of this document.

Sheet Dose_during extracts the examples for dose during exposure from sheet Import, and performs a check of the calculations used in the Monte Carlo Program.

Sheets Dose_life, Risk, and Haz_index perform similar checks for the lifetime average dose, risk estimate, and hazard index.

B.21 Dose_life_results.wb3

The Monte Carlo program (Appendix C) produces text files that contain the results. These text files are imported into this spreadsheet, and the results used to produce figures and numbers used in the remainder of this document. This spreadsheet sets up the information on lifetime average doses.

The Monte Carlo program produces a file temp.txt containing the results of the initial section of the main routine, in which the variability distribution for doses or the combined variability and uncertainty distribution for doses and risks are produced, (depending on the value of a switch in the program). The second section of the Monte Carlo program produces a file temp_unc.txt. These files are imported into the sheets Imp_false, Imp_true, Imp_time, and Imp_uncert in the following conditions:

Table B.2 Result files from the Monte Carlo program: imports to Risk_life_results.wb3	
Sheet of spreadsheet	**File imported, and conditions for Monte Carlo program**
Imp_false	temp.txt: 1,000,000 iterations of the first part of the program with the software switch with_uncert set to false (see Section C.6), and the remainder of the program set to include all uncertainties and variabilities.
Imp_true	temp.txt: 1,000,000 iterations of the first part of the program with the software switch with_uncert set to true (see Section C.6), and the remainder of the program set to include all uncertainties and variabilities.
Imp_time	temp.txt: 1,000,000 iterations of the first part of the program with the software switch with_uncert set to true (see Section C.6). Set Result=0 in the routine Tfisheater.extra_time to set the effective extra fishing duration (Section 6.3.2) to zero.
Imp_uncert	temp_unc.txtt: 50,000/5,000 iterations of the second part of the program with the software switch with_uncert set to true (see Section C.6), and the remainder of the program set to include all uncertainties and variabilities. This is the main set of results.
Imp_notime	temp_unc.txt: 10,000/1,000 iterations of the second part of the program with the software switch with_uncert set to true (see Section C.6). Comment out D_Time_eat.update_uncertainty in routine Tfisheater.update_uncertainty to remove the uncertainty in period of fish-eating and initial age (for the sensitivity analysis).
Imp_nomeal	temp_unc.txt: 10,000/1,000 iterations of the second part of the program with the software switch with_uncert set to true (see Section C.6). Comment out D_meals_per_year.update_uncertainty in routine Tfisheater.update_uncertainty to remove the uncertainty in number of meals per year (for the sensitivity analysis).
Imp_nocook	temp_unc.txt: 10,000/1,000 iterations of the second part of the program with the software switch with_uncert set to true (see Section C.6). Comment out cook_survival.update_uncertainty in routine Tfisheater.update_uncertainty to remove the uncertainty in PCB survival during cooking (for the sensitivity analysis).

Imp_noconc	temp_unc.txt: 10,000/1,000 iterations of the second part of the program with the software switch with_uncert set to true (see Section C.6). Comment out fishes[fish].update_uncertainty in routine TFish_Concs.update_uncertainty to remove the uncertainty in PCB concentrations in fish (for the sensitivity analysis).
Imp_notrend	temp_unc.txt: 10,000/1,000 iterations of the second part of the program with the software switch with_uncert set to true (see Section C.6). Comment out time_trend:=D_time_trend.random in routine TFish_Concs.update_uncertainty to remove the uncertainty in the time trend of PCB concentration in fish (for the sensitivity analysis).
Imp_all	temp_unc.txt: 10,000/1,000 iterations of the second part of the program with the software switch with_uncert set to true (see Section C.6). Comment out D_Time_eat.update_uncertainty D_meals_per_year.update_uncertainty fish_concs.update_uncertainty cook_survival.update_uncertainty in routine Tfisheater.update_uncertainty to remove all but numerical uncertainty (for the sensitivity analysis).
Imp_pop	temp_unc.txtt: 50,000/5,000 iterations of the second part of the program with the software switch with_uncert set to true (see Section C.6), and the remainder of the program set to include all uncertainties and variabilities, but with the alternate specification for the population uncertainty (see the routine Tfisheater.create).

Once these text files from the Monte Carlo program have been imported, the other sheets perform a few calculations and provide ready access to the distributions for exporting values (the spreadsheets were electronically linked with the master version of this document) and setting up graphs.

Sheet MLE_values shows the variability distribution for lifetime average dose (at the MLE value for uncertainty) in 0.1% steps, together with approximate estimates for uncertainties in the percentage point (±1 standard deviation).

Sheet Uncertainty contains the variability/uncertainty distribution for lifetime dose as 1% points on the uncertainty distributions for 31 selected percentiles of the variability distribution. Columns AM and AN contain estimates of cancers/year and total cancers, using the fixed U.S. EPA upper bound potency estimate of 2 kg-day/mg. Columns A and B export various values used in this document.

Sheet Combined shows the combined uncertainty/variability distribution for lifetime dose in 0.1% steps, together with approximate estimates for the uncertainties at each percentage point (±1 standard deviation), with columns G and H showing corresponding risk estimates using the fixed U.S. EPA upper bound potency of 2 kg-day/mg.

The Sensitivity sheet contains the results of sensitivity analyses, in which various uncertainties were omitted from the calculations.

There are multiple graphs defined in the spreadsheet to provide figures for this document.

B.22 Dose_while_results.wb3

This spreadsheet sets up the information on doses during exposure. The raw data is linked from the imported material in Dose_life_results.wb3.

Sheet MLE_values shows the variability distribution for average dose during exposure (at the MLE value for uncertainty) in 0.1% steps, together with approximate estimates for uncertainties in the percentage point (±1 standard deviation).

Sheet Uncertainty contains the variability/uncertainty distribution for dose during exposure as 1% points on the uncertainty distributions for 31 selected percentiles of the variability distribution.

Sheet Combined shows the combined uncertainty/variability distribution for dose during exposure in 0.1% steps, together with approximate estimates for the uncertainties at each percentage point (±1 standard deviation).

The Sensitivity sheet contains the results of sensitivity analyses, in which various uncertainties were omitted from the calculations.

There are multiple graphs defined in the spreadsheet to provide figures for this document.

B.23 Risk_results.wb3

This spreadsheet sets up the information on lifetime cancer risks. The raw data is mostly linked from the imported material in Dose_life_results.wb3. In addition, the file prob.txt produced by the Monte Carlo program is imported in sheet Probabilities to extract the probability for zero cancers for the combined variability/uncertainty distribution with potency uncertainty included or excluded (in the latter case, using the fixed U.S. EPA upper bound potency estimate of 2 kg-day/mg).

Sheet Uncertainty contains the variability/uncertainty distribution for cancer risk as 1% points on the uncertainty distributions for 31 selected percentiles of the variability distribution. Columns AM and AN contain estimates of cancers/year and total cancers. Columns A and B export various values used in this document.

Sheet Combined shows the combined uncertainty/variability distribution for lifetime cancer risk in 0.1% steps, together with approximate estimates for the uncertainties at each percentage point (±1 standard deviation). Columns F and G show corresponding risk estimates when the additional exposure time due to continuing body burden (Section 6.3.2) is omitted (obtained from the Monte Carlo program by always setting this time to zero).

Sheet Probabilities shows the probabilities for various numbers of cancers (from 0 to 100) as produced in the file probs.txt by the Monte Carlo program. Two sets of values are given, the first (column C) including the uncertainties in carcinogenic potency, the second (column D) excluding such uncertainties and using the fixed U.S. EPA upper bound potency of 2 kg-day/mg for all Aroclors. The long tail of the first distribution can be seen in the relatively constant, small, probability for large numbers of cancers — the table omits approximately 0.20% of the probability, but at least 16% of the contribution to the expected value.

There are multiple graphs defined in the spreadsheet to provide figures for this document.

B.24 HI_results.wb3

This spreadsheet sets up the information on hazard indexes. The raw data is linked from the imported material in Dose_life_results.wb3.

Sheet Uncertainty contains the variability/uncertainty distribution for hazard index as 1% points on the uncertainty distributions for 31 selected percentiles of the variability distribution.

Sheet Combined shows the combined uncertainty/variability distribution for hazard index in 0.1% steps, together with approximate estimates for the uncertainties at each percentage point (±1 standard deviation).

There are multiple graphs defined in the spreadsheet to provide figures for this document.

B.25 References for this appendix

Atkin, C. (1994). *A survey study of anglers residing near the Kalamazoo River basin.* Michigan State University.

Brown, J.F. Jr. (1994). Determination of PCB metabolic, excretion, and accumulation rates for use as indicators of biological response and relative risk. *Environ. Sci. Technol. 28*:2295–2305.

Brunner, S., Hornung, E., Santl, H., Wolff, E., Piringer, O.G., Altschuh, J., and Brüggemann, R. (1990). Henry's law constants for polychlorinated biphenyls: experimental determination and structure–property relationships. *Environ. Sci. Technol. 24*:1751–1754.

Frame, G. M., Cochran, J. W., and Boewadt, S.S., 1996. Complete PCB congener distributions for 17 Aroclor mixtures determined by 3 HRGC systems optimized for comprehensive, quantitative, congener-specific analysis. J. High Resol. Chromatogr., 19:657-668. Data available from http://www.epa.gov/toxteam/pcbid/down.htm.

Gilbert, R.O. (1987). Statistical methods for environmental pollution monitoring. Van Nostrand Reinhold, NY, NY.

Karasek, G.L.B (1998). *Hunting Results, Michigan. Small Game Seasons, 1992-1996.* Michigan Department of Natural Resources, Wildlife Division Report No. 3285. February, 1998. Obtainable from (at May 11, 2001) http://www.dnr.state.mi.us/pdfs/hunting/smallgame_9296results.pdf.

Land, C.E. (1971). Confidence Intervals for Linear Functions of the Normal Mean and Variance. *Ann. Math. Stat. 43*:1187–1205.

Land, C.E. (1973). Standard Confidence Limits for Linear Functions of the Normal Mean and Variance. *J. Amer. Statist. Assoc. 68*: 960–963.

Land, C.E. (1974). Confidence interval estimation for means after data transformations to normality. *J. Amer. Statist. Assoc. 69*:795–802.

Land, C.E. (1975). Tables of Confidence Limits for Linear Functions of the Normal Mean and Variance. *Selected Tables in Mathematical Statistics, Volume III*, 385–419.

Land, C.E. (1988). Hypothesis Tests and Interval Estimates," in *Lognormal Distributions, Theory and Applications*, E.L. Crow and K. Shimizu, eds. Marcel Dekker.

LTI (2000). Limno-Tech, Inc. Remedial Investigation Report and database of water measurements

Lyon, B.F., and C.E. Land (1999). *Computation of Confidence Limits for Linear Functions of the Normal Mean and Variance.* Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6285. ORNL/TM-1999/245. (Available at http://falcon.sis.utk.edu/ConfLimit/ at 9/6/2000).

MiCPHA (2000a). Klaviter, E., Humphrey, H., Bloomer, A.W., and Welch, R. *Kalamazoo River Angler Survey and Biological Testing Study, Final Report.* Submitted by Environmental Epidemiology Division, Community Public Health Agency, State of Michigan. Printed by ATSDR. May 2000.

MiCPHA (2000b). *Kalamazoo River Angler Survey and Biological Testing Study, Database for Phase I.* Provided by the Environmental Epidemiology Division, Community Public Health Agency, State of Michigan, in response to a Freedom of Information Request. Cover letter dated August 31, 2000, from Dr. Robert L. Wahl (Environmental Epidemiology Division) to Dr. Edmund Crouch (Cambridge Environmental Inc.).

MiCPHA (2000c). *Kalamazoo River Angler Survey and Biological Testing Study, Database for Phase II.* Provided by the Community Public Health Agency, State of Michigan, in response to a Freedom of Information Request. Cover letter dated December 23, 1998, from Dr. David R. Wade (Division of Environmental Epidemiology) to Ms. Dawn E. Penniman (Blasland, Bouck & Lee, Inc.).

Royston, J.P. (1982). Algorithm AS 181: The W test for normality. *Appl. Statist. 31*:176–180.

Royston, P. (1993). A toolkit for testing for non-normality in complete and censored samples. *Statistician 42*:37–43.

Royston, P. (1995). Remark AS R94. A remark on algorithm AS 181: The W-test for normality. *Appl. Statist. 44*:547–551.

Appendix C Details of the Monte Carlo analysis

The text describes all the distributions used in the analysis. Section C.1 gives details of the standard distributions used to build the distributions described in the text. Section C.2 summarizes the main programming objects used in the implementation. Section C.3 lists the electronic files for the Monte Carlo analysis available in the supplemental material (these may be used to completely reproduce the analysis). Section C.4 contains references for the previous sections. Sections C.5 and C.6 list the interface for the main program, and the main routine, respectively, for those who wish to browse such material (the complete program and all support routines are included in the additional material — see Section C.3)

C.1 Random number generation

The Monte Carlo technique strictly requires random numbers for its implementation. For our implementation, we used the standard computer technique of approximating this ideal with pseudo-random numbers. The basic generator is for the uniform distribution. All the other distributions are generated from the uniform distribution, so that given a uniform random number generator, the other distributions are generated exactly (Devroye, 1986).

C.1.1 Standard uniform pseudo-random variate generation

The generator used is a linear congruential generator of the form:

$$x_{n+1} = Ax_n + 1 \mod B$$

where:

x_n	is a sequence of 8-byte (64-bit) integers,
A	= 6,364,136,223,846,793,005 (hexadecimal 5851F42D4C957F2D), and
B	$= 2^{64}$

and all integer arithmetic is performed exactly. The multiplier A is an "excellent" one that passes the spectral test (Knuth, 1998).

A sequence of real values in the range [0,1) is obtained by multiplying the 8-byte (64-bit) integer x_n by 2^{-64}, and these real values are returned with full 64- bit precision by using the 80-bit extended precision real format of the INTEL® iAPX® coprocessors.

The seeds used for each run may be chosen arbitrarily by hand or approximately randomly from the system clock (in this application they were chosen using the system clock). Four word (2-byte) values are obtained, and adjoined to form the 8-byte seed. Seeds may be retrieved or set using either four word values or a single 8-byte integer value. Such retrieval and setting allows repeating exactly the analysis, or pieces of it, requiring solely the saving of a single 8-byte value at the location of the desired repeat (this technique was used to re-generate the examples).

During all calculations, intermediate real values are held with 64-bit precision, again using the 80 bit extended precision real data format. Results are stored with the 52-bit precision of the IEEE standard double precision real number, all conversions from higher precision being performed by rounding.

C.1.2 Arbitrary uniform random variates

Given a standard uniform random variate U, a random variate uniform on the range $[a,b)$ is obtained as:

$$X = a + (b-a)U$$

C.1.3 Triangular distribution random variates

The triangular distribution is defined by a range (a,b) and a point p where the distribution has its mode. Given a standard uniform random variate U, a random variate T from such a triangular distribution is obtained as:

$$\text{if } U < (p-a)/(b-a) \quad \text{then } T = a + \sqrt{(b-a)(p-a)U}$$
$$\text{else } T = b - \sqrt{(b-a)(b-p)(1-U)}$$

C.1.4 Piecewise linear random variates

The piecewise linear random variation is defined by a list of cumulative probabilities $0=P_0<P_1<P_2.<...<P_{N-1}=1$, and a corresponding list of values $V_0<V_1<V_2<V_2<..<V_{N-1}$. The cumulative distribution is defined by

$$P(x) = 0 \qquad \text{for } x \le V_0$$
$$P(x) = \frac{P_{i-1}(V_i - x) + P_i(x - V_{i-1})}{V_i - V_{i-1}} \qquad \text{for } V_{i-1} \le x \le V_i$$
$$P(x) = 1 \qquad V_{N-1} \le x$$

for $1 \le i \le N-1$. Given a standard uniform random variate U, a random variate L from this distribution is obtained by binary searching in the sequence $P_0,P_1,P_2.,...P_{N-1}$ for the index i such that $P_{i-1} \le U \le P_i$, and then

$$L = \frac{V_{i-1}(P_i - U) + V_i(U - P_{i-1})}{P_i - P_{i-1}}$$

C.1.5 Exponential random variates

Given a standard uniform random variate U, an exponential random variate E with density function $e^{-\lambda x}/\lambda$ is obtained as:

$$E = -\frac{\ln U}{\lambda}$$

C.1.6 Normal random variates

Given standard uniform variates U_1, U_2, two independent standard (mean zero, standard deviation unity) random normal variates are generated as:

$$N_1 = \sqrt{-2\ln U_1}\sin 2\pi U_2$$
$$N_2 = \sqrt{-2\ln U_1}\cos 2\pi U_2$$

In the implementation of this generator, either none or two variates are generated. If no unused variate is available, then two are generated — one is returned and the other stored. If an unused variate is available in store, it is returned immediately. This has the potential effect of introducing an undesired memory into the generator. The effect is minimized by flushing the store whenever the basic uniform pseudo-random generator's seed is set or read (see Section C.1.1).

Given a standard normal random variate N, a variate n from a normal distribution with mean μ and standard deviation σ is generated as:

$$n = \mu + N\sigma$$

C.1.7 Truncated normal variates

Truncated normal variates are obtained by repeatedly generating a normal random variate with the required mean and standard deviation until the resulting value is within the truncation range, returning the first such normal variate satisfying that condition.

C.1.8 Lognormal random variates

Given a standard normal random variate N, a lognormal variate with median $\exp(\mu)$ and geometric standard deviation σ is generated as:

$$L = \exp(\mu + N\sigma)$$

C.1.9 Gamma random variates

A $\Gamma(a,b)$ random variate has density function:

$$\frac{b^{-a}x^{a-1}e^{-x/b}}{\Gamma(a)}$$

In an obvious pseudo-code (where `random` returns a standard uniform variate), the algorithm for a gamma random variate is (Devroye, 1986):

```
c:=a-ln(4); λ:=Sqrt(2*a-1); d:=a+λ;
if (a<1) then begin  { Johnk's generator }
      repeat
            y:=random^(1/a); z:=random^(1/(1-a));
      until (y+z) ≤ 1;
      return -ln(random)*y/(y+z);
end else begin   { Cheng's rejection algorithm GB }
      repeat
            u:=random; v:=random;
            y:=ln(v/(1-v))/λ; x:=a*exp(y); z:=u*Sqr(v); r:=c+d*y-x;
      until r ≥ ln(z);
      return b*x;
end
```

C.1.10 Mean of a lognormal distribution

An uncertainty distribution for the mean of a lognormal distribution, based on a sample of size N, has been obtained (Land, 1971, 1973, 1974, 1975, 1988; Lyon & Land, 1999). This distribution has not been evaluated in any form that allows ready sampling from it. However, for a sample from a normal distribution, S^2/σ^2 is chi-squared distributed with N–1 degrees of freedom, where σ is the (unknown) true standard deviation, S^2 is the usual unbiased estimate of standard deviation, and$(M-\mu)\sqrt{N}/\sigma$ is independently normally distributed, where μ is the true mean and M is the sample mean. This suggests the following algorithm as an approximation:

> Take a sample χ^2 from a chisquared (N-1) distribution and compute $s^2 = S^2(N\text{-}1)/\chi^2$.
> Take a sample z from a normal (0,1) distribution and compute $m = M+sz/\sqrt{N}$.
> Return $\exp(m+s^2/2)$.

This algorithm has been tested against the distribution computed by Lyon & Land (1999) — see Appendix B.6. It is found to be biased high at the left end, approaching unbiased for large numbers of degrees of freedom and large standard deviation. As such, it was considered adequate for this assessment.

C.1.11 Multinormal random variate

A d-dimensional vector Y of multinormal random variates with ($d \times d$ dimensional) variance-covariance matrix Σ is obtained by providing the lower diagonal matrix H such that $HH^T=\Sigma$. H necessarily exists since Σ is symmetric and positive definite. Let $X=(N_1, N_2, N_3, \ldots N_d)^T$ where the N_i are independent standard (zero mean, unit variance) normal variates. Then $Y=HX$ is the required vector.

C.2 The implementation

The Monte Carlo analysis was implemented in Object Oriented Pascal using Borland Delphi 5.0 (from Inprise Corporation). For the fish ingestion scenario, the classes (objects) defined were:

TTimeEating	Handles the distribution of lifetime period eating fish, and the initial age of fish eating.
TMealsPeryear	Handles the distribution of meals per year, conditional on time spent eating.
TConcInfo	Handles the details of summary statistics for concentrations measured in a single species of fish on a single occasion in a single ABSA.
TAllConcs	Contains a linked list of TConcInfo in order to contain information on fish concentrations for a single species in a single ABSA over multiple years.
TOneFishConcs	Contains an array of TAllConcs to handle a single species of fish in all the ABSAs.
TFishConcs	Contains an array of TOneFishConcs to handle all species of fish.
TFishFracs	Handles the empirical data on the fractions of meals that are of various fish species.
TCookSurvival	Handles the distributions for PCB survival during cooking by various methods.
TPCB_data	Data on PCB toxicity values, and the fraction of PCBs in various fish.
TFishEater	Represents a person by containing a TCookSurvival, TFishFracs, TFishConcs, TMealsPeryear, and a TTimeEating.

Each of these classes has methods that allow rapid Monte Carlo analysis. The TFishEater, for example, has Update_Variability and Update_Uncertainty methods that invoke the similar

methods for the relevant other classes contained within it. The interface for these classes is shown in Section C.5, and the main routine that uses them is shown in Section C.6. The complete listing of the program and all support files is not shown here since they occupy several thousand lines (much of the code in the support files is not used), but the complete code in electronic format accompanies this report to allow replication. The interface and main routine (below) show the outline of the methodology.

To achieve reasonable numerical stability where the variability/uncertainty distributions were required, the variability loop was iterated 50,000 times, and the uncertainty loop was iterated 5,000 times, effectively sampling a total population of 250,000,000 fish-eaters. For simpler cases, where just the variability was required (for doses), or where both variability and uncertainty were combined (for the random individual), 1,000,000 iterations were performed. Each variability loop (50,000 iterations) took about 1.1 seconds on a 500 MHz Pentium III machine with 64 Mbyte RAM, so a complete uncertainty/variability computation took about 90 minutes (various compiler settings affect this time, and smaller amounts of RAM are also likely to substantially increase it). With 1,000,000 iterations, the percentiles of the dose variability distribution are numerically stable to about 2 significant figures at the 99th percentile (and better for less extreme percentiles). With 1,000,000 iterations, the percentiles of the combined variability/uncertainty distributions (for a random individual) are numerically stable to about ±1 digit in the second significant figure at the 99th percentile. With the 50,000/5,000 iterations for the variability/uncertainty calculations, the values computed at the 95th percentile of the 95th percentile are numerically stable to better than 20%.

C.3 Supplemental information for the Monte Carlo analysis

Supplemental material in electronic form accompanies this document in a Zipped archive file. That material includes the following files:

readme.txt	This documentation in ASCII format
Kalamazoo_F.dpr	Delphi project file This just contains a compiler directive to run as a console application, and a main program that executes `mainroutine`.
Kala_fish.pas	All the Kalamazoo-specific material
StringFunctions.pas	Support file
ConvertUnits.pas	Support file (not used, but required for compilation)
Distributions.pas	Support file
incbetafunction.pas	Support file
MathLib.pas	Support file
MonteCarlo.pas	Support file
Multi_Distributions.pas	Support file

Normals.pas	Support file
Randoms.pas	Support file
SortUnit.pas	Support file
ClassRegistry.pas	Support file (not used, but needed for compilation)
fish_PCB.dat	Data file (see Kala_fish.pas)
fish_data.dat	Data file (see Kala_fish.pas)
eat_data.dat	Date file (see Kala_fish.pas)

These data files are used to initialize some of the objects in Kala_fish.pas. They need to be placed in a directory that is specified in Kala_fish.pas by the constant data_root_dir. Output will go to the same directory, in files called temp.txt and temp_unc.txt. Read `mainroutine` in Kala_fish.pas to find the formats for all these files.

The results of the Monte Carlo program are produced in data files temp.txt, temp_unc.txt, probs.txt, and examples.txt. These results (in some cases from multiple runs of the program under differing conditions) have been imported (using the Tools/DataTools/QuickColumns/Parse command) into the following Quattro Pro spreadsheets that are part of the supplementary material and are discussed in Appendix B. Some additional calculations are performed in these spreadsheets, and graphs are set up there.

dose_life_results.wb3	Results for lifetime average doses (see Appendix B.21)
dose_while_results.wb3	Results for doses during exposure (see Appendix B.22
HI_results.wb3	Results for hazard index (see Appendix B.24)
Risk_results.wb3	Results for lifetime risk (see Appendix B.23)
Examples.wb3	Examples (see Appendix B.20)

C.4 References for this appendix

Knuth, D.E., 1998. The art of computer programming, Vol. 2: Seminumerical Algorithms, Third Edition. Addison-Wesley Publishing Co., Reading, MA. (ISBN 0-201-89684-2)

Devroye, L., 1986. Non-uniform random variate generation. Springer-Verlag, NY. (ISBN 0-387-96305-7 and 3-540-96305-7). Errata and addenda available at http://jeff.cs.mcgill.ca/~luc/ .

Land, C.E. (1971). Confidence Intervals for Linear Functions of the Normal Mean and Variance. *Ann. Math. Stat. 43*:1187–1205.

Land, C.E. (1973). Standard Confidence Limits for Linear Functions of the Normal Mean and Variance. *J. Amer. Statist. Assoc. 68*: 960–963.

Land, C.E. (1974). Confidence interval estimation for means after data transformations to normality. *J. Amer. Statist. Assoc. 69*:795–802.

Land, C.E. (1975). Tables of Confidence Limits for Linear Functions of the Normal Mean and Variance. *Selected Tables in Mathematical Statistics, Volume III*, 385–419.

Land, C.E. (1988). Hypothesis Tests and Interval Estimates," in *Lognormal Distributions, Theory and Applications*, E.L. Crow and K. Shimizu, eds. Marcel Dekker.

Lyon, B.F., and C.E. Land (1999). *Computation of Confidence Limits for Linear Functions of the Normal Mean and Variance.* Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6285. ORNL/TM-1999/245. (Available at http://falcon.sis.utk.edu/ConfLimit/ at 9/6/2000).

C.5 Code interface for major classes

Here is the code for the main interface section. Many other support files are used, but examination of the interface indicates the methodology adopted. The following types are provided by support files not listed here, but provided in the electronic accompanying files:

TDistribution	A class that has descendants defined to represent many of the standard distributions.
PVector	A pointer to a linear array of arbitrary size (about 2Gbyte maximum).
TMultiNormalDistribution	An implementation of the multinormal distribution (not yet integrated with TDistribution).

```
{ ************************************************************ }
{ ********************** } Interface { ********************** }
{ ************************************************************ }
type
  fish_types=(walleye,sucker,carp,bass,pike,panfish,catfish,turtle);
  Aroclor=(A1016,A1242,A1248,A1254,A1260);
  ABSA_range=3..9;  { Valid ABSAs for this analysis }
  ABSA_list=array[fish_types] of ABSA_range;
                        { A list of ABSAs, one for each fish type }
  meal_fracs=array[fish_types] of double;
                        { The fraction of meals of a particular fish type }

const
  data_root_dir='d:\project\b-1287 PCBs in Kalamazoo\progdata\';
  fish_data_file=data_root_dir+'fish_data.dat';
                                { Fish concentration information }
{}
{ Format: lines of text with the following form: }
{ Carp        93        3        11        1.491        0.504      8.95}
```

```
{ name        year      absa     Number    mean log     sd of log  maxval}
{ name = fish name; must be identical to one of fish_name_string (below) }
{ year = years since 1900 for sampling }
{ absa = ABSA number, should be in ABSA_range }
{ Number = number of samples for this fish type in this year in this ABSA }
{ mean log = mean of logs of total PCB concentration in mg/kg in those samples }
{ sd of log = unbiased estimate of sd of log total PCB conc in mg/kg }
{ maxval = maximum concentration in mg/kg seen in this fish in this ABSA }
{          in this year}
{ File can have any number of entries. }
{}
  eat_data_file=data_root_dir+'eat_data.dat';   { Meal fractions, by fish }
{}
{ Format: lines of text, 1 per person in the survey, with the following form }
{ First line:  Walleye Sucker Carp Bass Pike Panfish Catfish  Turtle  Total
{ 2nd - nth:   0.0000 0.0000 0.0000 0.6153  0.0000  0.3076  0.0769  0.0000  325
{  First line is names of fish for which next lines list meal fractions.  The }
{  initial names on the line must correspond exactly in spelling and order with }
{  the names in fish_names (below).  Any subsequent entries are ignored. }
{  So the heading "Total" and the column below it are ignored. }
{  Subsequent lines.  A list of meal-fractions for individuals in the highest }
{  quartile of meals-per-year (the "total" entry is actually meals per year in }
{  the data file used here), straight from the survey information.
{  After the highest quartile there is a blank line, followed by a similar }
{ list of lines for the next highest quartile, and so on. }
{}
  fish_pcb_file=data_root_dir+'fish_pcb.dat';  { Aroclor fractions, by fish ]
{}
{ Format: lines of text }
{ Line 1: }
{Aroclor    1016    1242    1248    1254    1260 }
{ The word "Aroclor" can be replaced with anything, but must be a single }
{ non-zero-length string with no blanks. }
{ List of names must agree in order and name with Aroclor_name_string below }
{ lines 2 onwards: example }
{ Carp       0.057   0.078   0.317   0.474   0.074  }
{ fish_name  1016_frac 1242_frac 1248_frac 1254_frac 1260_frac }
{ fish_name must agree exactly with one of those in fish_name_string, below }
{ If fish_name recurs, it overwrites the previous one.  All fish_names must }
{ occur. }
{ Fractions are interpreted as the fractions of the aroclor in line 1 that }
{ they match in order. }

{ Standard parameters }
  std_lifetime = 70;        { years }
  std_bodyweight = 70;      { kg }
  index_year=99.0;          { index year from which we start }
{ Data information }
  Max_eat_quartile=70;      { Maximum number of entries per quartile }
                            { in eat_data_file }
  fish_names:array[fish_types] of string =
        ('Walleye','Sucker','Carp','Bass','Pike','Panfish','Catfish','Turtle');
  fish_name_string=
        'Walleye Sucker  Carp    Bass    Pike    Panfish Catfish Turtle';
  Aroclor_name_string='1016 1242 1248 1254 1260';

  { Imposed limitation }
  Max_Val_Multiplier=10;
            { Average concentration cannot exceed Max_Val_Multiplier }
            { times the maximum concentration in any ABSA }


type
{ Define a class type to handle the evaluation of the time spent eating }
```

```
{ fish, and the initial age }
  TTimeEating = Class (TObject)
    tot_time,init_age:double;   { Calculated values for total and initial time }
    constructor create;
    procedure update_uncertainty;
       { Updates the parameters according to the uncertainty distribution}
    procedure update_variability;
     { Updates the tot_time and init_age according to variability distribution }
    function pop_entry_rate:double;
     { Return estimate of fraction of observed sample number entering per year }
    destructor destroy; override;
  private
{ Parameters of the variability distribution for total time }
    lam_1,lam_2,alpha:double; { Current values for total time. }
                              { Note that alpha is converted to the total time }
                              { distribution, not the survey time distribution }
{ Variability distribution }
    D_init_age:TDistribution;  { Initial age }
    P_init_age:PVector;        { Used for D_init_age }
{ Uncertainty distribution }
    D_uncert:TMultiNormalDistribution;  { Uncertainty }
    P_work:PVector;            { Work vector for uncertainty distribution }
  end;

{ Define a class to handle the number of meals per year }

  TMeals_per_year = Class (TObject)
    constructor create;
    destructor destroy; override;
    function random_meals(const t_eat:double):double;
     { Return random sample of # of meals per year from variability distribution }
     { given the length of time eating }
    procedure update_uncertainty;
  private
    D_uncert:TMultiNormalDistribution;  { Uncertainty }
    P_parms:PVector;      { The current parameters }
                          { mu_a=P_parms^[0], mu_b 1, sigma_a 2, sigma_b 3 }
  end;


{ TConc_info contains information on one fish type in one ABSA at one time. }
{ Designed to go in a linked list of these for multiple times. }
  TConc_info=class (TObject)
    mean,sd,maxval,  { Mean, SD of natural log of sample concentrations }
                     { and maximum measured value.  Concs in mg/kg. }
    offset:double;   { Years BEFORE index year }
    num:integer;     { Number of samples }
    next:TConc_info; { link to next one }
    constructor create(const m,s,yrsbefore:double; const n:integer);
      { mean, sd, years before index year, number of samples }
    constructor createfromstring(const s:string; const index:double);
      { Extract the required information from a line of text }
      { index is the index year, with the same basis as in the input string }
      { Expected format is }
      { fishtype year ABSA number mean sd }
      { This routine ignores fishtype and ABSA }
    function MeanContribution(const time_slope:double):double;
      { Contribution to mean in index year.  ln(mg/kg). }
      { argument is time decay factor (positive) }
    function VarContribution(const m,time_slope:double):double;
      { Contribution to variance in index year.  Square of ln(mg/kg). }
      { arguments are grand mean at index year, time decay factor (positive) }
    function MaxContribution:double;
      { Largest measured value, in mg/kg }
```

```
    procedure writeconcs(var f:Textfile); { Used for checking }
  end;

{ TAll_concs contains information on all times for one fish type in one }
{ ABSA }
  TAll_concs=class (TObject)
    mean,sd,maxval:double;
       { Current mean & standard deviation estimates (of log conc) in }
       { index year and maximum value in mg/kg at any time}
    num:integer;                { Number of measurements }
    constructor create;
    destructor destroy; override;
    procedure addconcs(const s:string; const index:double);
             { Add more information }
    procedure update_time_slope(const time_slope:double);
             { Adjust information to index year }
    procedure update_uncertainty;
       { Sample from distribution for mean concentration in index year }
       { Use update_time_slope just before this to use different time trend }
    procedure writeconcs(var f:Textfile);  { Used for checking }
  private
    meas_mean,meas_sd:double; { Mean & SD of measurements, adjusted to index year }
    Concs:TConc_info;         { Linked list of years of concentrations }
    D_Mean,D_sd:TDistribution; { Distributions for mean & SD construction }
    P_Mean,P_sd:PVector;       { Used for initializing D_mean & D_sd }
  end;


{ TOne_Fish_concs contains all the information for one fish in all ABSAs }
  TOne_Fish_concs = class (TObject)
    constructor create;
    destructor destroy; override;
    procedure update_time_slope(const time_slope:double);
             {Re-adjust to index year }
    procedure update_uncertainty;
             { Sample from distribution for mean in index year }
             { Needs update_time_slope just before this. }
    procedure addconcs(const s:string; const index:double);
             { add another fish entry }
    function ConcIn(const absa:ABSA_range):double;
      { Get the concentration for this fish in the given ABSA.  It is a fatal }
      { error if there are no data there. }
    function FindABSA(const absa:ABSA_range):ABSA_range;
      { Find the "nearest" ABSA for which there are data, in a random }
      { direction. }
    procedure writeconcs(var f:Textfile);  { used for checking }
  private
    all_concs:array[ABSA_range] of TAll_concs;
             { Data structure holding concentrations }
  end;


{ TFish_Concs contains information on all fish in all ABSAs }
{ No variability here; these are independent of people }

  TFish_Concs = Class (TObject)
    time_trend: double;           { Current estimate of time trend in concs. }
    constructor create(const s:string);  { s is a data file for fish concs }
    destructor destroy; override;
    function ConcIn(const fish:fish_types; const absa:ABSA_range):double;
      { Return the concentration in that type of fish in that ABSA.  It is a }
      { fatal error for that ABSA to have no information on that fish. }
    procedure FindABSA(const absa:ABSA_range; var a_list:ABSA_list);
      { Obtain a list of the "nearest" ABSAs to absa containing information }
```

```
      { on the given fish }
    procedure update_uncertainty;
    procedure writeconcs(var f:Textfile); { used for checking }
  private
    fishes:array[fish_types] of TOne_Fish_concs;  { Array holding fish conc }
                                                  { data for all fish & ABSAs }
    D_time_trend:TDistribution;   { Time trend distribution }
    P_time_trend:PVector;         { Used for D_time_trend }
  end;

{ TFish_fracs is a class to handle the fractions of meals that are a given }
{ fish }

  TFish_fracs = Class (TObject)
    q_size:array[1..4] of integer;  { Actual number in each quartile }
    eat_data:array[1..4,0..Max_eat_quartile-1] of meal_fracs;
    constructor create(const s:string); { s is a data file for fractions }
    function randommeal(const m_p_yr:double):meal_fracs;
  end;

{ Cooking survival }
  TCook_Survival = Class (TObject)
    cook_surv:double;  { current value from variability }
    constructor create;
    destructor destroy;override;
    procedure update_uncertainty;
    procedure update_variability;
  private
    D_bake,D_broil,D_fry:TDistribution;
    t_bake_surv,t_broil_surv,t_fry_surv:double;  { current values from uncertainty }
  end;

{ PCB data}
  TPCB_data = Class (TObject)
    constructor create(const s:string);
       { s is the name of the file containing the PCB fractions in fish }
       { See fish_pcb_file above }
    destructor destroy;override;
    function PCB_frac(const fish:fish_types; const aro:Aroclor):double;
       { return the fraction of Aroclor aro in the given fish }
    function potency(const aro:Aroclor):double;
    function RfD_Standard(const aro:Aroclor):double;
      { Value of RfD for standard duration of exposure, here 17.5 yrs }
    function RfD_time_factor(const t_dur:double):double;
      { Modifying value for RfD for other exposure periods }
    procedure update_uncertainty;
    procedure update_variability;
    procedure setreturnvalue(RetVal:TRandom);
  private
    PCB_pot_ratio:array[Aroclor] of double;  { Relative potencies }
    PCB_fractions:array[fish_types,Aroclor] of double;
                              { Aroclor fractions in fishes }
    D_pot_var,                { Potency variability distribution }
    D_pot_unc,                { Potency uncertainty distribution }
    D_RfD_var,                { RfD variability distribution }
    D_RfD_unc: TDistribution; { RfD uncertainty distribution }
    curr_pot_var,             { Current potency variability distribution value }
    curr_pot_unc,             { Current potency uncertainty distribution value }
    curr_RfD_var,             { Current RfD variability distribution value }
    curr_RfD_unc: double;     { Current RfD uncertainty distribution value }
    DReturnvalue:Trandom;     { Return fixed or variable results for RfD & Potency}
  end;

{ This will be the basic object on which we operate }
```

```
  TFishEater = Class (TObject)
    time_eating,              { time eating fish in a lifetime }
    init_age,                 { initial age of eating fish }
    t_extra,                  { extra time due to body burden }
    meals_per_year,           { average meals per year }
    fish_meal_mass,           { average fish mass per meal, in kg }
    PCB_cook_survival,        { survival of PCBs through cooking }
    Population                { Number of fishers }
            : double;
    meal_frac:meal_fracs;     { Avg. fraction for each fish type }
    curr_absa:ABSA_range;     { Current ABSA }
    curr_list:ABSA_list;      { Current ABSAs -- the "nearest" ABSA to curr_absa }
                              { that has data for each fish }
    D_Time_eat:TTimeEating;   { Object to calculate time_eating and init_age }
    D_Meals_per_year:TMeals_per_year;  { Object to calculate meals per year }
    fish_concs:TFish_Concs;   { All the relevant fish concs }
    fish_fracs:TFish_fracs;   { All the fish fractions of meals }
    Cook_survival:TCook_survival; { Survival through cooking }
    PCB_data:TPCB_data;       { Data on PCBs }
{ basic methods }
    constructor create;
    Destructor destroy; override;
{ Uncertainty and variability updates }
    procedure update_uncertainty;
    procedure update_variability;
{ Methods used in the update methods }
    function get_fish_meal_mass:double;  { Get the required random value }
    function get_curr_absa:ABSA_range;   { Get the current ABSA }
    function extra_time(const t_init,t_dur:double):double;
{ Method that returns the stuff of interest }
    function average_dose(var c_avg,dose_while,can_risk,haz:double):double;
      { Returns lifetime average dose in ug/kg-d, }
      { c_avg            Average concentration in fish }
      { dose_while      Average dose rate during exposure (ug/kg-d) }
      { can_risk        Cancer risk }
      { haz             Hazard index }
      { for current realization of uncertainty and variability (no fixed }
      { values) }
    procedure writeinputs(var ofile:TextFile; const conc,risk:double);
      { Write a list of the current input values used for the calculations in }
      { average_dose.  For convenience, we pass in the calculated }
      { average PCB concentration and relevant risk value. }
    procedure setreturnvalue(RetVal:TRandom);
      { Used to force RfD and Potency to fixed values }
  private
    D_population:TDistribution;
    function cum_eff_time(const t_dur:double):double;  { Integral of body burden }
  end;


{ ************************ END OF INTERFACE ************************* }
```

C.6 Main routine used

Here is the main routine used to generate the variability and uncertainty distributions. The TMonteCarlo class implements an array to hold results from a Monte Carlo simulation, to sort them, to calculate means and standard deviations, and to output percentiles and uncertainty estimates on those percentiles. Output was written to ASCII files, and imported to spreadsheets for plotting and further analysis.

```
{ ********************* MAIN ROUTINE *************************** }
procedure mainroutine;
{ The main routine }

type
  acc_type=(d_acc,l_acc,c_acc,h_acc);
   { d_acc     accumulator for dose during consumption }
   { l_acc     accumulator for lifetime average dose }
   { c_acc     accumulator for cancer risk }
   { h_acc     accumulator for hazard index }

const
{ Percentage points of variability distribution to accumulate for }
{ uncertainty analysis }
  num_ppts=31;
  ppoint:array[0..num_ppts-1] of double = (
  0.005,0.01,0.02,0.03,0.04,0.05,0.075,
  0.1,0.15,0.2,0.25,0.3,0.35,0.4,0.45,0.5,
  0.55,0.6,0.65,0.7,0.75,0.8,0.85,0.9,
  0.925,0.95,0.96,0.97,0.98,0.99,0.995);
{ Number of loops for variability and uncertainty analysis }
{ For the first section of code, 1,000,000 is sufficient. }
{ For the second section, use 50,000 & 5,000 for final run. }
{ For quickie runs for sensitivitiy analysis, use 10,000 and 1,000. }
  var_loops=50000;
  unc_loops=5000;
  outfile_var=data_root_dir+'temp.txt';
  outfile_unc=data_root_dir+'temp_unc.txt';
{ Examples: where, how many, and where to save them }
  ex_number=10;              { Number of examples in each set }
  ex_sets=2;                 { Number of example sets }
  ex_value:array[0..ex_sets-1,acc_type] of double=((0.05,0.05,1e-5,1.0),
                                                   (0.5,0.5,1e-4,10.0));
                             { Locations of examples in sets }
  outfile_exs=data_root_dir+'examples.txt';
{ Probabilities for given number of cancers }
  n_canc_probs=100;
  outfile_prob=data_root_dir+'probs.txt';
{ EPA potency, for calculating total cancers using fixed value }
  EPA_potency=2; { kg-d/mg }



var
  fisher:TFishEater;                    { what we work with }
  dose:array[acc_type,0..num_ppts+2] of TMonteCarlo;
          { Used for uncertainty distributions. }
          { 0..num_ppts hold percentage points of variability distribution. }
          { Entry num_ppts holds the mean }
          { Entry num_ppts+1 holds popn. product }
          { Entry num_ppts+2 holds integrated popn. product }
  var_dose:array[acc_type] of TMonteCarlo;
```

```
                { Used for variability distributions }
  var_dose_i:array[acc_type] of TMonteCarlo_with_index;
                { Used to allow repeat of variability+uncertainty calcs for }
                { selected entries. }
  i,j,k:integer;                              { Counters }
  acc:acc_type;                               { Counter }
  pc,blo,bhi,mean,sd:double;                  { Temporaries }
  ofile:TextFile;                             { Output file }
  l_dose,c_risk,h_risk:double;                { More temporaries }
  w_dose,c_avg:double;                        { More temporaries }
  randkey:int64;                              { A saved random seed }
  ex_first:integer;                           { Location of first example }
  ex_sort:acc_type;                           { What do we want examples for? }
  new_risk:double;                            { Example calculated risk }
  with_uncert:boolean;
  ex_num:integer;                             { checked number of examples }
  ex_set_no:integer;                          { Set of examples }
  prob_num:array[acc_type,0..n_canc_probs] of double;{ Probabilities for # of cancers
}
  t_prob:double;


begin
{ ************************ }
{ ***** First Section ***** }
{ ***** Combine uncertainty and variability  *********** }
{ ***** (or omit uncertainty for dose calc.) *********** }
{ ************************ }

  with_uncert:=true;    { Toggle for including or excluding uncertainty }
                        { in first section.  Omitting it gives useful results }
                        { for dose only.  Including it gives the standard }
                        { combined uncertainty/variability distributions -- the }
                        { uncertainty distributions for a randomly chosen }
                        { individual. }
{ Create our object }
  fisher:=TFishEater.create;  { Sets up fisher at the MLE for uncertainty }
                              { for doses. }

  if (with_uncert) then begin
     fisher.setreturnvalue(randomval);
  end
  else begin
     fisher.setreturnvalue(fixedval);
  end;
{ Create the objects to hold results of variability loops }
  for acc:=low(acc_type) to high(acc_type) do begin
     var_dose_i[acc]:=TMonteCarlo_with_index.create(var_loops,-1);
  end;
{ Do an MLE or all together run }
  for i:=1 to var_loops do begin
    randkey:=getrandseed64;
    if (with_uncert) then fisher.update_uncertainty;
    fisher.update_variability;
    l_dose:=fisher.average_dose(c_avg,w_dose,c_risk,h_risk);
    var_dose_i[d_acc].addin(w_dose,randkey);  { Dose while exposed }
    var_dose_i[l_acc].addin(l_dose,randkey);  { Lifetime average dose }
    var_dose_i[c_acc].addin(c_risk,randkey);  { Cancer risk }
    var_dose_i[h_acc].addin(h_risk,randkey);  { Hazard index }
    if((i mod 1000)=0) then write('.');  { Keep informed of progress }
  end;
  writeln(' sorting MLE');
  for acc:=low(acc_type) to high(acc_type) do var_dose_i[acc].sort;
  writeln(' Output MLE ');
```

```
  Assignfile(ofile,outfile_var);
  Rewrite(ofile);
  for i:=1 to 999 do begin
    for acc:=low(acc_type) to high(acc_type) do begin
      pc:=var_dose_i[acc].percentile(i/1000,blo,bhi);
      write(ofile,pc:12,' ',blo:12,' ',bhi:12,'  ');
    end;
    writeln(ofile);
  end;
{ Output means and SD for straight distribution }
  for acc:=low(acc_type) to high(acc_type) do begin
    mean:=var_dose_i[acc].mean(sd);
    write(ofile,mean:12,' ',sd:12,' ',0.0:12,'  ');  { last zero keeps format }
  end;
  writeln(ofile);
{ Output logarithmic mean and sd }
  for acc:=low(acc_type) to high(acc_type) do begin
    mean:=var_dose_i[acc].genmean(sd,hide_log);
    write(ofile,mean:12,' ',sd:12,' ',0.0:12,'  ');  { last zero keeps format }
  end;
  writeln(ofile);
  CloseFile(ofile);
{ Now write out some examples }
  writeln('Now producing examples');
  Assignfile(ofile,outfile_exs);
  rewrite(ofile);
  for ex_set_no:=0 to ex_sets-1 do begin
{ Examine all available distributions.  }
    for ex_sort:=low(acc_type) to high(acc_type) do begin
      if (ex_number>var_loops) then ex_num:=var_loops else ex_num:=ex_number;
  { Locate position to output examples }
      var_dose_i[ex_sort].findvalue(ex_value[ex_set_no,ex_sort],ex_first);
  { Write header }
      case ex_sort of
         d_acc: write(ofile,'Dose during distribution at
',ex_value[ex_set_no,ex_sort]:5:3,' ug/kg-d');
         l_acc: write(ofile,'Lifetime dose distribution at
',ex_value[ex_set_no,ex_sort]:5:3,' ug/kg-d');
         c_acc: write(ofile,'Cancer risk distribution at
',ex_value[ex_set_no,ex_sort]:11);
         h_acc: write(ofile,'Hazard index distribution at
',ex_value[ex_set_no,ex_sort]:5:2);
      end;
      writeln(ofile,' at the ',100*ex_first/var_loops:6:2,' %ile');
      fisher.writeinputs(ofile,0,0);  { writes a header }
  { Set up the location of values to be output }
      ex_first:=ex_first-(ex_num div 2);
  { Check for within bounds }
      if (ex_first+ex_num-1)>=var_loops then ex_first:=var_loops-ex_num;
      if (ex_first<0) then ex_first:=0;
  { Compute }
      for i:=ex_first to ex_first+ex_number-1 do begin
        setrandseed64(var_dose_i[ex_sort].index(i));
        if (with_uncert) then fisher.update_uncertainty;
        fisher.update_variability;
        l_dose:=fisher.average_dose(c_avg,w_dose,c_risk,h_risk);
  { Check that we have reproduced exactly the previous results }
        new_risk:=0;
        case ex_sort of
           d_acc: new_risk:=w_dose;
           l_acc: new_risk:=l_dose;
           c_acc: new_risk:=c_risk;
           h_acc: new_risk:=h_risk;
        end;
```

```
        if (new_risk<>var_dose_i[ex_sort].entry(i)) then begin
           writeln('Does not check at entry ',i,'  ',new_risk,'
',var_dose_i[ex_sort].entry(i));
        end;
  { Write out the current values of inputs saved in fisher, together with }
  { the average PCB concentration in fish and the result }
        fisher.writeinputs(ofile,c_avg,new_risk);
      end;
    end;
  end;
  close(ofile);
{ Clean up }
  for acc:=low(acc_type) to high(acc_type) do begin
    var_dose_i[acc].free;
  end;
  writeln('Done first section.  Hit Enter for second section, ^C to abort.');
  readln;

{ ************************************** }
{ ********* Second Section *************** }
{ ********* Variability and uncertainty treated separately to *********** }
{ ********* obtain the 2-dimensional distributions            *********** }
{ ************************************** }

  writeln('Uncertainty analysis');
{ Ensure everything is random }
  fisher.setreturnvalue(randomval);
{ Create variability holders, and make some more holders }
  for acc:=low(acc_type) to high(acc_type) do begin
    var_dose[acc]:=TMonteCarlo.create(var_loops,-1);
    for k:=0 to num_ppts+2 do dose[acc,k]:=TMonteCarlo.create(unc_loops,-1);
  end;
{ Zero our probability array }
  for acc:=low(acc_type) to high(acc_type) do begin
    for k:=0 to n_canc_probs do prob_num[acc,k]:=0;
  end;
{ Uncertainty loop }
  for j:=1 to unc_loops do begin
    fisher.update_uncertainty;
{ Variability loop }
    for i:=1 to var_loops do begin
      fisher.update_variability;
      l_dose:=fisher.average_dose(c_avg,w_dose,c_risk,h_risk);
      var_dose[d_acc].addin(w_dose);  { Dose during exposure }
      var_dose[l_acc].addin(l_dose);  { Lifetime average dose }
      var_dose[c_acc].addin(c_risk);  { Cancer risk }
      var_dose[h_acc].addin(h_risk);  { Hazard index }
    end;
{ Extract the information on percentiles }
    for acc:=low(acc_type) to high(acc_type) do begin
      var_dose[acc].sort;
      for k:=0 to num_ppts-1 do begin
         dose[acc,k].addin(var_dose[acc].percentile(ppoint[k],blo,bhi));
      end;
{ Get the mean over the variability distribution = population mean }
      mean:=var_dose[acc].mean(sd);
      dose[acc,num_ppts].addin(mean);
{ Next statements are meaningless for d_acc and h_acc, but get the }
{ cancers/yr for l_acc and c_acc }
      mean:=mean*fisher.D_time_eat.pop_entry_rate*fisher.Population;  { mean dose *
popn entry rate }
{ For the lifetime dose rate, convert to mg/kg-d and use EPA potency }
{ For c_acc, the potency has already been applied }
      if (acc=l_acc) then mean:=(mean/1000)*EPA_potency;
```

```
      dose[acc,num_ppts+1].addin(mean);
{ Then get total number ever (again, just for l_acc and c_acc }
{ Meaningless for d_acc and h_acc }
      mean:=mean/fisher.fish_concs.time_trend; {mean dose * popn entry rate/time_trend
}
      dose[acc,num_ppts+2].addin(mean);
{ Got all we want, so clear for next variability loop }
      var_dose[acc].clear;
    end;
{ Accumulate probabilities for given numbers of cancers }
    for acc:=low(acc_type) to high(acc_type) do begin
{ Get last entered value }
      mean:=dose[acc,num_ppts+2].lastadded;
      case acc of
        d_acc,h_acc: { do nothing -- meaningless entries here } ;
        l_acc,c_acc:
          begin
            t_prob:=exp(-mean);
            prob_num[acc,0]:=prob_num[acc,0]+t_prob;
            for k:=1 to n_canc_probs do begin
              t_prob:=t_prob*mean/k;
              prob_num[acc,k]:=prob_num[acc,k]+t_prob;
            end;
          end;
      end;
    end;
    if ((j mod 10)=0) then write('.');
  end;
{ Order the uncertainty distribution for percentiles and the mean }
  writeln('Sorting uncertainty');
  for acc:=low(acc_type) to high(acc_type) do begin
     for k:=0 to num_ppts+2 do dose[acc,k].sort;
  end;
  writeln('Writing results');
{ Write out all these results }
  Assignfile(ofile,outfile_unc);
  Rewrite(ofile);
  for acc:=low(acc_type) to high(acc_type) do begin
    for k:=0 to num_ppts-1 do begin
      write(ofile,ppoint[k]:12,',');
    end;
    writeln(ofile,'mean,peryr,alltime');
    for i:=1 to 99 do begin
      for k:=0 to num_ppts+2 do begin
        write(ofile,dose[acc,k].percentile(i/100,blo,bhi):12,',');
      end;
      writeln(ofile);
    end;
{ The following is inefficient, but it works OK and saves having to define }
{ yet more data structures }
{ Write out means and standard deviations at fixed percentage points }
    for k:=0 to num_ppts+2 do begin
      mean:=dose[acc,k].mean(sd);
      write(ofile,mean:12,',');
    end;
    writeln(ofile);
    for k:=0 to num_ppts+2 do begin
      mean:=dose[acc,k].mean(sd);
      write(ofile,sd:12,',');
    end;
    writeln(ofile);
{ Now do the same for logarthms of results }
    for k:=0 to num_ppts+2 do begin
      mean:=dose[acc,k].genmean(sd,hide_log);
```

```
      write(ofile,mean:12,',');
    end;
    writeln(ofile);
    for k:=0 to num_ppts+2 do begin
      mean:=dose[acc,k].genmean(sd,hide_log);
      write(ofile,sd:12,',');
    end;
    writeln(ofile);
    writeln(ofile,'spacer line,');
  end;
  CloseFile(ofile);
{ Output probabilities }
  Assignfile(ofile,outfile_prob);
  Rewrite(ofile);
  for k:=0 to n_canc_probs do begin
    writeln(ofile,k,',',                                    { Number of cancers }
                  prob_num[c_acc,k]/unc_loops:13,',',       { Prob. with potency
uncertainties }
                  prob_num[l_acc,k]/unc_loops:13);          { Prob. with EPA potency }
  end;
  CloseFile(ofile);
{ Cleanup for uncertainty }
  for acc:=low(acc_type) to high(acc_type) do begin
    for k:=0 to num_ppts+2 do dose[acc,k].free;
  end;
  for acc:=low(acc_type) to high(acc_type) do begin
    var_dose[acc].free;
  end;
{ End of uncertainty piece }

  fisher.free;
end;


{ ******************  END OF MAIN ROUTINE ******************* }
```

SDMS US EPA Region V

Imagery Insert Form

Some images in this document may be illegible or unavailable in SDMS.
Please see reason(s) indicated below:

☐ Illegible due to bad source documents. Image(s) in SDMS is equivalent to hard copy.

Specify Type of Document(s) / Comments:

☐ Includes ___ COLOR or RESOLUTION variations.
Unless otherwise noted, these pages are available in monochrome. The source document page(s) is more legible than the images. The original document is available for viewing at the Superfund Records Center.

Specify Type of Document(s) / Comments:

☐ Confidential Business Information (CBI).
This document contains highly sensitive information. Due to confidentiality, materials with such information are not available in SDMS. You may contact the EPA Superfund Records Manager if you wish to view this document.

Specify Type of Document(s) / Comments:

☒ Unscannable Material:
Oversized __ __ or _X__ Format.
Due to certain scanning equipment capability limitations, the document page(s) is not available in SDMS. .

Specify Type of Document(s) / Comments:

OTHER DATA ON CD

☐ Document is available at the EPA Region 5 Records Center.

Specify Type of Document(s) / Comments: